의학박사 학위논문

종양미세환경 내 유전자 발현 분석을 통한 PD-L1 발현과 면역 회피 조절 기전 분석

The comprehensive transcriptomic analysis regarding the mechanisms of PD-L1 expression and immune modulation in tumor microenvironment

2018년 2월

서울대학교 대학원
의학과 중개의학전공
옥 찬 영

A thesis of the Doctor of Philosophy degree

The comprehensive transcriptomic analysis regarding the mechanisms of PD-L1 expression and immune modulation in tumor microenvironment

종양미세환경 내 유전자 발현 분석을 통한 PD-L1 발현과 면역 회피 조절 기전 분석

February 2018

Department of Medicine
Translational Medicine Major
Seoul National University College of Medicine
Chan-Young Ock

Abstract

Introduction: PD-L1 expression, existence of tumor-infiltrating lymphocyte (TIL), and mutational burden would be promising biomarkers to predict the efficacy of anti-PD-1/PD-L1 inhibitor, but clinical implication of these markers have not been completely established. In the current study, we analyze the immunogenomic properties in The Cancer Genome Atlas (TCGA) according to the classification of tumor based on PD-L1 expression and TIL. Moreover, upstream mechanism of PD-L1 expression and immune-suppressive signature of TIL were comprehensively analyzed.

Methods: The RNA expression levels of PD-L1 and CD8A in the samples in the pan-cancer database of TCGA (N=9,677) were analyzed. Based on their median values, the samples were classified into four tumor microenvironment immune types (TMITs). In tumor tissues of head and neck squamous cell carcinoma (HNSCC) of Seoul National University Hospital (SNUH cohort), PD-L1 expression and epithelial-mesenchymal transition (EMT) markers were assessed by immunohistochemistry and this finding was validated in 5 HNSCC cell lines.

Results: In TCGA analysis, the number of somatic mutations, PD-L1 amplification and oncogenic virus infection were significantly associated with TMIT I, respectively. Immune-suppressive signature by regulatory T-cells or M2 macrophage clearly affected survival among TME with highly infiltrated CD8-positive T-cells. PD-L1 expression was associated with EMT phenomenon, assessed by high vimentin and low E-cadherin expressions. Increased PD-L1 expression with EMT-

positive was significantly associated with poor prognosis in SNUH cohort and TCGA cohort. PD-L1 expression was increased by cisplatin treatment in SNUH cohort and HNSCC cell lines, accompanying MAPK/ERK pathway.

Conclusions: This integrative analysis highlights the importance of the assessment of both PD-L1 expression and TIL signature in the era of immune checkpoint inhibitors.

Keywords: Tumor Microenvironment; Comprehensive analysis; Next-generation sequencing; Transcriptome; Immune evasion; Immune checkpoint inhibitor

Student Number: 2014 – 22036

CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

CTLA-4; cytotoxic T-lymphocyte Associated Protein 4

PD-1; Programmed Cell Death 1

PD-L1; Programmed Cell Death 1 Ligand

Introduction

A recent strategy targeting immune checkpoints such as CTLA-4 and PD-1/PD-L1 shows promising clinical benefits and introduces a paradigm shift in cancer treatment. Immune checkpoint-blocking agents have shown a remarkable clinical efficacy with a long response duration in immunogenic tumors (1-5).

Although the expression of PD-L1 on the surface of tumor cells, as measured by immunohistochemistry, may potentially serve as a predictive factor to identify patients who would benefit from anti-PD-1/PD-L1 therapy, not all PD-L1-positive patients respond well (4-6). Interestingly, the degree of TIL infiltration and PD-L1 expression in the tumor microenvironment (TME) are also correlated with the clinical outcomes of anti-PD-1/PD-L1 therapies (6, 7). Moreover, recent advances in immuno-genomics have shown that tumors with a high mutational burden, abundant neoantigen and microsatellite (MSI)-high status are associated with a good response to anti-PD-1/PD-L1 therapy (8-14). Moreover, oncogenic viruses such as Epstein-Barr virus (EBV) or human papillomavirus (HPV) are associated with an inflamed TME, and a favorable clinical outcome in response to anti-PD-1/PD-L1 therapy would also be expected (10, 15). Because of the complex nature of tumor immunity, a comprehensive immuno-genomic analysis that include analyses of the interaction between the tumor and TIL in

the TME and investigations of the underlying reasons for the promotion of tumor immunogenicity is needed.

The classification of tumors into four different types based on the presence or absence of CD8+ CTLs and PD-L1 expression was recently suggested (15, 16). Tumors with high PD-L1 expression and the presence of CD8+ CTLs in the microenvironment are classified as TME immune type I, which would only benefit from anti-PD-L1/PD-1 therapies (15, 16). However, this concept has not been investigated with a large-scale genomics database because a standardized methodology for assessing PD-L1 and CD8+ CTLs has not been clarified.

Moreover, the upstream biologic mechanisms of PD-L1 expression in cancer has been poorly clarified, although possible pathways have been suggested. Firstly, Viruses and epithelial-mesenchymal transition (EMT) are associated with high PD-L1 expression (17). PD-L1 up-regulation occurs in cancers associated with human papillomavirus (HPV) such as uterine, cervical, head and neck cancers (18-23), possibly by interferon-gamma secreted from tumor-infiltrating immune cells via the JAK/STAT pathway (20-24). Moreover, PD-L1 expression is also associated with the mesenchymal signature of tumors. For example, more than half of sarcoma and sarcomatoid lung carcinoma patients showed high PD-L1 expression levels, irrespective of tumor type (25, 26). EMT changes, manifested by E-cadherin (encoded by *CDH1*) down-regulation and vimentin (encoded by *VIM*)

up-regulation, also correlated with PD-L1 induction (27). HNSCC is a suitable model to investigate the clinicopathologic features associated with PD-L1 up-regulation. Recent advances in genomics have shown that mutational HPV-positive and -negative tumor profiles, which have retained and lost p16 expression, respectively, clearly differ (28-30).

Secondly, PD-L1 expression would be changed by chemotherapy. Increasing *in vitro* evidence has shown that the expression of PD-L1 in tumors changes in response to various exogenous signals including interferon-gamma, radiotherapy, or chemotherapeutic agents (31-33). Therefore, PD-L1 changes during the treatment course may be also evident in HNSCC, but few studies have examined this.

In the current study, we classified a large set of TCGA pan-cancer samples into four tumor microenvironment immune types (TMIT) by measuring the mRNA expression levels of PD-L1 and CD8A with in-depth analysis focused on immune-suppressive microenvironment. Moreover, we analyzed the association of PD-L1 expression and EMT status, and the change of PD-L1 expression by cisplatin treatment, which is the most commonly used chemotherapeutic agent to treat HNSCC.

Material and Methods

Processing of genomic data from The Cancer Genome Atlas project

We used publicly available, level 3 data of TCGA in the current study. Clinical information, gene-level somatic mutation data, copy number variation (CNV) data, and mRNA expression data obtained by RNA sequencing (RNAseq) of the TCGA samples were downloaded from the UCSC Cancer Browser (https://genome-cancer.ucsc.edu) on June 3, 2015. To count the number of total somatic mutations, multiple somatic mutations including non-synonymous mutations, insertion-deletion mutations, and silent mutations were each counted and summated, and germline mutations without somatic mutations were excluded. The amplification and deletion statuses in the CNV thresholded data, which was calculated using Gistic 2.0 (34), were documented as “2” and "-2", respectively, and mRNA expression data, which were generated using the Illumina HiSeq V2 platform, are presented as reads per kilobase per million (RPKM) and transformed into log 2 values for analysis. The patients’ clinical data and microsatellite instability (MSI) status of the tumors were extracted from the TCGA database. The MSI status was available for 1,164 samples, which included COAD (N = 426) and STAD samples (N = 414). The

detection of oncogenic viruses, such as Epstein-Barr virus (EBV), human papillomavirus (HPV), and hepatitis B virus (HBV), in each tumor sample (N = 6,511) and the neoantigen number (N = 3,726) were also referenced in a previous report by Rooney et al. (10). Altogether, samples of 32 cancer types (N = 10,354) were included in the analysis.

Statistical analyses of genomic data

According to previous reports regarding the four types of immune evasion of cancer (15), after merging log 2-transformed values of the RPKM of PD-L1 and CD8A, we divided all of the TCGA samples into four groups as follows: Type I, PD-L1 expression higher than the median and CD8A expression higher than the median; Type II, PD-L1 expression lower than the median and CD8A expression lower than the median; Type III, PD-L1 expression higher than the median and CD8A expression lower than the median; and Type IV, PD-L1 expression lower than the median and CD8A expression higher than the median. The cytolytic activity of each sample was calculated using the log 2-transformed value of the geometric mean of GZMA and PRF1, as previously reported (10). The prognostic significance of the four tumor microenvironment immune types (TMITs) was estimated using Kaplan-Meier plots (log-rank test) and Cox proportional hazards regression analysis. The statistical significance of two continuous variables, such as the total number of somatic mutations and the number of

neoantigens, as well as the CD8A expression level, cytolytic activity and PD-1 expression was calculated by linear regression analysis. The significance of the differences between continuous values, such as the number of mutations, interferon-gamma expression, and cytolytic activity, and categorical variables, such as TMIT, MSI status, and POLE mutation, was calculated by the Wilcoxon rank sum test or analysis of variance with Tukey's post hoc test (for comparisons of more than 3 groups). Comparing the proportion of each TMIT according to categorical variables was performed by Fisher's exact test. The number of somatic mutations and copy number variation, such as the amplification and deletion of a gene in a previously suggested set of 373 genes that are frequently altered in cancer (35) with PD-L1 (CD274), were analyzed by logistic regression to determine whether these variables can significantly predict a specific TMIT. Univariate and multivariate logistic regression analyses were performed to determine whether there is a significant association between clinicopathologic characteristics and the ability to predict TMIT I. To analyze immune-suppressive microenvironment, we calculated infiltration of CD8-positive T-cells, regulatory T-cells, and M2 macrophages using Cibersort algorism (36). Among tumor samples whose CD8-positive T-cells infiltration score is higher than median, regulatory T-cells or M2 macrophage infiltration status was classified by highly deviated (+/- 0.5 standard deviation) to fitted line of

regulatory T-cells score or M2 macrophage score by CD8-positive T-cells infiltration score, respectively.

Statistical analyses of microarray, RNA sequencing, and clinicopathologic data

The BRB-ArrayTools software program (http://linus.nci.nih.gov/BRB-ArrayTools.html) was used in the analysis of gene expression data (37). A heatmap was generated using Cluster and TreeView software programs (38). To assess EMT characteristics of tumors from an HNSCC cohort of TCGA study (28) and cell lines from CCLE (39), 75 gene expression signatures associated with EMT from a previous study, GSE4824 (40), and a previously developed approach were used (41, 42). Briefly, the 75 gene expression data in the training set (GSE35640, Supplementary Table 1) were combined to form a classifier according to a Bayesian compound covariate predictor (BCCP) (43). The robustness of the classifier was assessed using a misclassification rate determined during leave-one-out cross-validation in the training set. The BCCP classifier estimated the likelihood that an individual patient had either an epithelial or mesenchymal signature according to a Bayesian probability cut-off of 0.5, which was optimized by comparing results with a previously reported proportion of EMT signatures in HNSCC (44). A comparison

of continuous values, such as H-scores for E-cadherin and vimentin, or the log 2 value of reads per kilobase of transcript per million mapped reads (RPKM) of PD-L1, was undertaken using a Wilcoxon Rank Sum test. A Fisher's exact test was used to determine associations between clinicopathologic parameters. The significance of clinicopathologic factors on PD-L1-positivity was calculated by logistic regression. OS was measured from the diagnosis date until death, or the last follow-up date if censored. Progression-free survival (PFS) was calculated from the first day of definitive treatment up to the date of disease progression, confirmed by imaging, death, or the last follow-up date if censored. Survival analyses were carried out according to the Kaplan-Meier method with log-rank testing to assess differences between groups. A Cox proportional hazard regression model was used for univariate and multivariate survival analyses. The RPKM cut-off for PD-L1 in survival analyses was determined by a median value of PD-L1 RPKM for mesenchymal signatures.

All reported P values were two-sided, and considered significant if $P < 0.05$. A false discovery rate was applied to control type I errors. All statistical analyses and data generation were carried out using R version 3.1.3 (http://www.r-project.org) and STATA version 12 (StataCorp LP, College Station, TX, USA).

HNSCC patient cohort

Medical records were retrospectively reviewed to identify patients diagnosed with HNSCC who were treated at Seoul National University Hospital (SNUH) from December 2004 to November 2012. Patients from whom paraffin-embedded tumor samples were obtained prior to and after cisplatin chemotherapy were included.

Cisplatin treatment in HNSCC patients

The decision to treat patients was determined by a multidisciplinary team. Patients were treated initially with induction chemotherapy and/or definitive CCRT or radical surgery (including primary tumor and regional lymph node dissection). Induction chemotherapy regimens included docetaxel, cisplatin, or 5-fluorouracil. CCRT regimens consisted of cisplatin. Radiotherapy delivered 5 days per week using a simultaneous integrated boost technique. Gross tumor lesions or high-risk volumes received 63 Gy to 67.5 Gy in 28 to 30 fractions over 6 weeks using a daily dose of 2.25 Gy, and low- and intermediate-risk volumes were irradiated to 48 Gy to 56 Gy using a daily dose of 1.8 Gy to 2.0 Gy with concurrent chemotherapy of weekly cisplatin.

Immunohistochemistry

Representative, formalin-fixed, paraffin-embedded tissue blocks from each case were submitted for immunohistochemistry (IHC) using the following antibodies: mouse anti-p16 (E6H4) monoclonal antibody (mAb; Roche/MTM/Ventana Medical Systems, Tucson, AZ, USA), mouse anti-E-cadherin (36B5) mAb (Novocastra Laboratories, Newcastle upon Tyne, UK), mouse anti-vimentin (V9) mAb (Dako, Ely, UK) and rabbit anti-PD-L1 (E1L3N) XP® mAb (Cell Signaling Technology, Danvers, MA, USA). IHC was performed using the Ventana Benchmark XT system (Ventana Medical Systems). When tissue sections showed diffuse and strong nuclear and cytoplasmic staining in ≥ 70% of tumor cells, this was considered positive for p16 (45). For E-cadherin and vimentin, staining intensity was scored in four categories: no staining (0), weak (1+), moderate (2+) and strong (3+) staining. The percentage of tumor cells showing the different staining intensities were evaluated by a trained pathologist. An IHC score (H-score) was then calculated using the following formula: 1 × (percentage of cells showing weak staining) + 2 × (percentage of cells showing moderate staining) + 3 × (percentage of cells showing strong staining). E-cadherin and vimentin were used as an epithelial or mesenchymal phenotype marker, respectively. An EMT phenotype was defined as

low E-cadherin expression with an H-score < 200 and high vimentin expression with a H-score > 30. PD-L1 IHC was evaluated based on the intensity and proportion of membranous staining, with or without cytoplasmic staining, in tumor cells and was scored as follows: 0, less than 5% of tumor cells; 1, weak in ≥ 5% of tumor cells; 2, moderate in ≥ 5% of tumor cells; and 3, strong in ≥ 5% of tumor cells. Cases showing membranous staining for PD-L1 in ≥ 5% of tumor cells (i.e., including IHC scores 1, 2 or 3) were considered PD-L1-positive.

Head and neck cancer cell lines experiments

Head and neck cancer cell lines were purchased from the American Type Culture Collection (Manassas, VA, USA) and Korean Cell Line Bank (Seoul, Korea) and cultured as previously described (46). Briefly, the SNU-1066, SNU-1041 and SNU-1076 cell lines were maintained in RPMI 1640 medium containing 100 U/mL penicillin, 100 μg/mL streptomycin (Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (GIBCO, Grand Island, NY, USA). The Detroit-562, and FaDu and PCI-13 cell lines were maintained in American Type Culture Collection Eagle's modified essential medium (EMEM) with 100 U/mL penicillin, 100 μg/mL streptomycin (Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (GIBCO). All cell lines were incubated under standard culture conditions (5% CO_2 at 37°C). Cells were resuspended

in lysis buffer (Cell Signaling Technology), incubated on ice for 10 min, and centrifuged for 15 min at 4°C. Samples containing equal quantities of total protein were resolved on SDS–polyacrylamide denaturing gels, transferred to polyvinylidene fluoride membranes, and probed with antibodies according to the manufacturer's protocols. Antibodies against PD-L1, p-MEK Ser217/221, MEK, p-STAT3, STAT3, E-cadherin, vimentin, and β-actin were purchased from Cell Signaling Technology. β-Actin was used as the protein loading control Detection was performed using an enhanced Lumi-Light Western Blotting Substrate kit (Roche, Basel, Switzerland).

Flow cytometry analysis for cell lines

Flow cytometry was performed as previously described (47). A total of 2×10^5 cells was aliquoted and placed into assay tubes. Next, 2 mL of fluorescence-activated cell sorting (FACS) buffer was added to each tube and rinsed twice by centrifugation. The cells were resuspended in 100 mL of FACS buffer with fixable viability dye (eBioscience, San Diego, CA). The cells were stained with PD-L1 phycoerythrin (PE; eBioscience) or isotype control for 30 min on ice in staining buffer (2% bovine serum albumin and 0.01% sodium azide). Analysis was conducted using a FACSCalibur instrument (BD Biosciences, Franklin Lakes, NJ, USA) with CELLQuest software (BD

Biosciences). Flow cytometric analysis of PD-L1 expression in head and neck cancer cells was conducted. After 24 h, the cells were harvested and stained with either mouse anti-human PD-L1 (clone 5H1) or a mouse IgG1 isotype control followed by PE-conjugated goat anti-mouse Ig.

Ethics

This study was approved by the Institutional Review Board of Seoul National University Hospital (approval number: H-1307-051-504) and was conducted in accordance with the Principles of the Declaration of Helsinki.

Results

Distribution and clinical implication of TMIT across cancer types.

We analyzed 9,677 tumor samples from 32 cancer types included in the TCGA dataset. The median of the log 2-transformed RPKM values, which indicate the mRNA expression level, of PD-L1 and CD8A were 4.72 and 6.97, respectively. The log transformation has no effect on the division at the median of the distributions, but only improves the analysis of the correlation. The expression levels of PD-L1 and CD8A were generally positively correlated even though a considerable proportion of tumor samples were found to be PD-L1-high and CD8A-low or PD-L1-low and CD8A-high tumors ($P < 0.001$, $R^2 = 0.250$, **Fig. 1a**). CD8A expression was significantly correlated with cytolytic activity (10) and PD-1 expression ($P < 0.001$, $R^2 = 0.718$ and 0.712, respectively).

All tumor samples were divided into four groups of TMIT according to the median values of PD-L1 and CD8A expression. Among all of the evaluated samples, 34.6% of the samples were classified as TMIT I, defined by high PD-L1 expression in the tumor and surrounding TME and high CD8A expression, which indicates a high proportion of CD8+ CTLs. The proportions of TMIT II (low PD-

L1/low CD8A), III (high PD-L1/low CD8A), and IV (low PD-L1/high CD8A) were 34.6%, 15.4%, and 15.4%, respectively. The proportion of TMIT I, II, III, and IV samples of skin cutaneous melanoma (SKCM) were 43.4%, 33.2%, 5.9%, and 18.4%, respectively, and these values are comparable to those included in previous reports (38%, 41%, 1%, and 20%, respectively) (15, 16).

The proportion of TMIT samples was analyzed according to cancer type (**Fig. 1b**). As expected, cancers derived from lymphoproliferative tissues, such as thymoma (THYM) and diffuse large B cell lymphoma (DLBC), had the highest proportion of TMIT I tumors among all cancer types (84.0% and 70.8%, respectively), further supporting the hypothesis that the signature reliably reflects the TMIT in cancer tissues. Among the solid cancers, lung adenocarcinoma (LUAD, 67.1%), kidney clear cell carcinoma (KIRC, 64.8%), lung squamous cell carcinoma (LUSC, 63.5%), and head and neck squamous cell carcinoma (HNSC, 54.1%) had the highest proportion of TMIT I samples. The clinicopathological features of the TCGA patients are summarized as follows: older patients (> 60 years) had a significantly higher proportion of TMIT I tumors compared with younger patients, although the actual difference between two groups was small (37.4% versus 32.4%, $P < 0.001$).

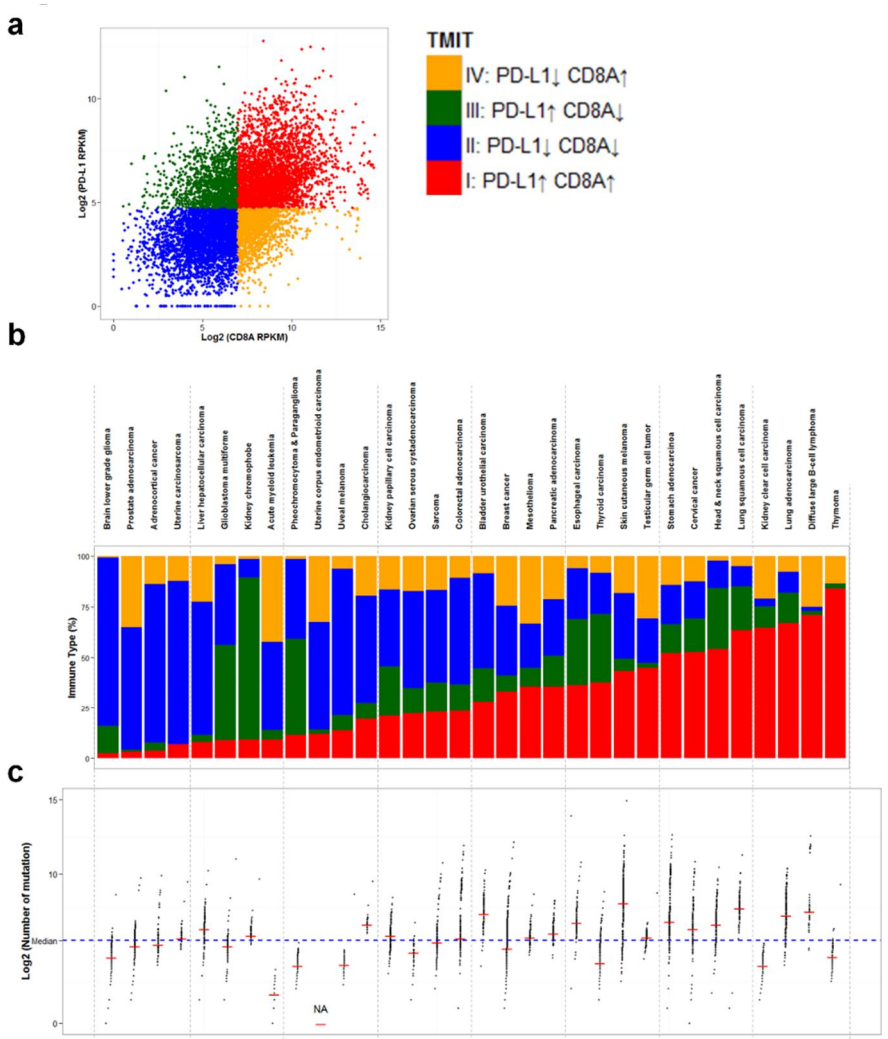

Figure 1. Distribution of tumor microenvironment immune type and mutational burden across cancer type. A scatter plot of log 2-transformed values of reads per kilobase of transcript per million reads mapped (RPKM) of PD-L1 and CD8A is shown (a). The portion of tumor microenvironment immune types (TMIT, I, red [34.6%]; II, blue [34.6%]; III, green [15.4%]; IV; yellow [15.4%]) (b) and the log 2-transformed value of the total number of somatic mutations according to cancer types of The Cancer Genome Atlas (c) are graphed. The red bar indicates the median value of each column, and the blue dashed line indicates the median value of the total samples. Abbreviation: NA, not

High mutational burden is associated with TMIT I

Because previous reports have shown that the degree of mutational burden and the presence of neoantigen are correlated with the immunogenic features of the tumor and reliably predict a good response to an anti-PD-1/PD-L1 treatment strategy (10, 12-14), we compared the proportion of different TMIT groups according to the mutational burden. According to cancer type, we observed a tendency toward a correlation between the total number of somatic mutations (hereby, number of mutations) as well as the number of neoantigens and the proportion of TMIT I (**Fig. 1c**). Interestingly, TMIT I tumors had a significantly higher number of mutations as well as a higher number of neoantigens compared with the other TMIT groups ($P < 0.001$, **Fig. 2a**). Tumor samples with a higher number of mutations than the median value (46) had a significantly higher proportion of TMIT I compared with those with fewer mutations (40.3% versus 30.4%, $P < 0.001$, **Fig. 2b**).

The microsatellite instability (MSI) has also been reported as a type of high mutational burden with immunogenic features (14, 48). As expected, MSI-high tumors had a significantly higher number of mutations compared with MSI-low tumors and microsatellite stable (MSS) tumors (**Fig. 2c**). Moreover, MSI-high tumors had a high proportion of TMIT I compared MSI-low tumors and MSS tumors (52.2% versus 25.0% and 29.9%, respectively, $P < 0.001$, **Fig. 2d**).

Because the mutational burden is significantly associated with TMIT I, we then assessed the association of TMIT I with somatic mutations in 373 genes that are frequently mutated in many cancers (35) as well as the *PD-L1* (*CD274*) gene (**Table 1)**. We found that *VHL*, *PBRM1*, *CASP8*, and *ATM* mutations are significantly associated with TMIT I. In contrast, *EGFR* and *BRAF* mutations were found to be correlated with TMIT III.

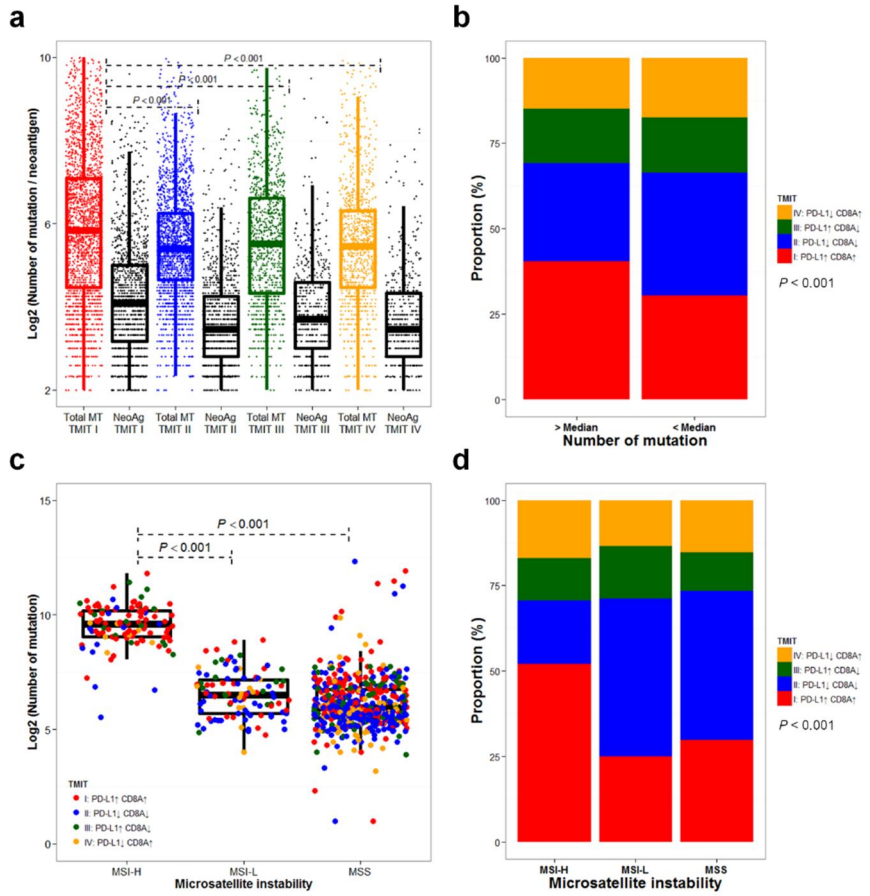

Figure 2. High mutational burden is associated with tumor microenvironment immune type I. Box plot of log 2-transformed values of the number of total somatic mutations (first, third, fifth, and seventh columns, the color scheme is as follows: I, red; II, blue; III, green; IV, yellow) and neoantigens (second, fourth, sixth, and eighth columns) according to tumor environment immune types (TMIT) are plotted (a). Proportion of TMIT according to a number of total mutations higher or lower than the median is compared (b). Box plot of log 2-transformed value of the number of mutations according to microsatellite instability (MSI) status is plotted (c). The proportion of TMIT according to MSI status is compared (d). Abbreviations: Total MT, total number of somatic mutations; NeoAg, number of neoantigens; MSI-H, microsatellite instability-high; MSI-L, microsatellite instability-low; MSS, microsatellite stable; RPKM, reads per kilobase of transcript per million reads mapped.

Table 1. Specific genomic alterations of cancer driver genes to predict tumor microenvironment immune type I.

Gene	Genomic alteration	Odds ratio	*P* value	FDR
VHL	mutation	3.37	2.21×10^{-20}	4.15×10^{-18}
CASP8	mutation	3.13	2.56×10^{-8}	8.81×10^{-7}
PBRM1	mutation	2.40	6.59×10^{-12}	8.23×10^{-10}
NCOR1	mutation	2.37	2.58×10^{-8}	8.81×10^{-7}
MXRA5	mutation	2.13	5.93×10^{-10}	4.44×10^{-8}
FAT1	mutation	1.99	1.44×10^{-8}	6.00×10^{-7}
ANK3	mutation	1.95	1.20×10^{-8}	5.64×10^{-7}
MUC17	mutation	1.94	1.90×10^{-11}	1.78×10^{-9}
BRAF	mutation	1.70	5.50×10^{-9}	3.44×10^{-7}
FLG	mutation	1.69	7.05×10^{-9}	3.78×10^{-7}
CD274	amplification	3.46	2.05×10^{-14}	7.43×10^{-12}
APC	amplification	3.33	6.10×10^{-9}	3.68×10^{-7}
DIAPH1	amplification	3.09	1.16×10^{-9}	1.40×10^{-7}
CDX1	amplification	2.73	1.69×10^{-8}	8.75×10^{-7}
NPM1	amplification	2.54	2.27×10^{-9}	1.65×10^{-7}
PIGZ	amplification	1.67	4.94×10^{-10}	8.94×10^{-8}
MUC4	amplification	1.64	2.03×10^{-9}	1.65×10^{-7}
ACVR2B	deletion	4.50	1.54×10^{-11}	5.57×10^{-9}
MYD88	deletion	4.32	3.56×10^{-11}	5.58×10^{-9}
TGFBR2	deletion	3.42	4.64×10^{-11}	5.58×10^{-9}
RPSA	deletion	4.26	1.52×10^{-10}	1.31×10^{-8}
SLC22A14	deletion	3.92	1.81×10^{-10}	1.31×10^{-8}
EAF1	deletion	4.10	6.79×10^{-10}	3.50×10^{-8}
MRPS25	deletion	4.19	6.45×10^{-10}	3.50×10^{-8}
CTNNB1	deletion	3.50	7.86×10^{-10}	3.55×10^{-8}
ZNF620	deletion	3.78	1.55×10^{-9}	6.21×10^{-8}

Odds ratio was calculated that the proportion of TMIT I in each genomic alteration-positive divided by those in the genomic alteration-negative, by logistic regression.

Filtering condition of this table is odds ratio > 1.5 and FDR < 1.00×10^{-7}.

Abbreviation: FDR, false discovery rate; TMIT, tumor microenvironment immune type.

PD-L1 amplification is associated with TMIT I

The association of *PD-L1* amplification and TMIT was analyzed because *PD-L1* amplification has been reported to serve as a good predictive marker of the response to anti-PD-1/PD-L1 therapy in Hodgkin's lymphoma (4, 18) and has been found to be associated with high immune cytolytic activity (10). The frequency of *PD-L1* amplification in all cancers was 1.6% (149 out of 9,364 samples), and HNSC (4.8%), sarcoma (4.7%), ovarian serous cystadenocarcinoma (4.3%), and DLBC (4.2%) had a relatively high frequency of *PD-L1* amplification (**Fig. 3a**). Compared with the wild-type tumors, a high proportion of tumors with *PD-L1* amplification were TMIT I (64.4% versus 34.4%, $P < 0.001$, **Figs. 3b** and **3c**).

In addition to *PD-L1* amplification, the analysis of 373 oncogenic genes in cancer revealed that *APC*, *NPM1*, and *CDX1* amplifications as well as *CTNNB1* deletion are associated with TMIT I. Moreover, *EGFR* amplification and *CDKN2A* deletion were found to be associated with TMIT III (**Table 1**).

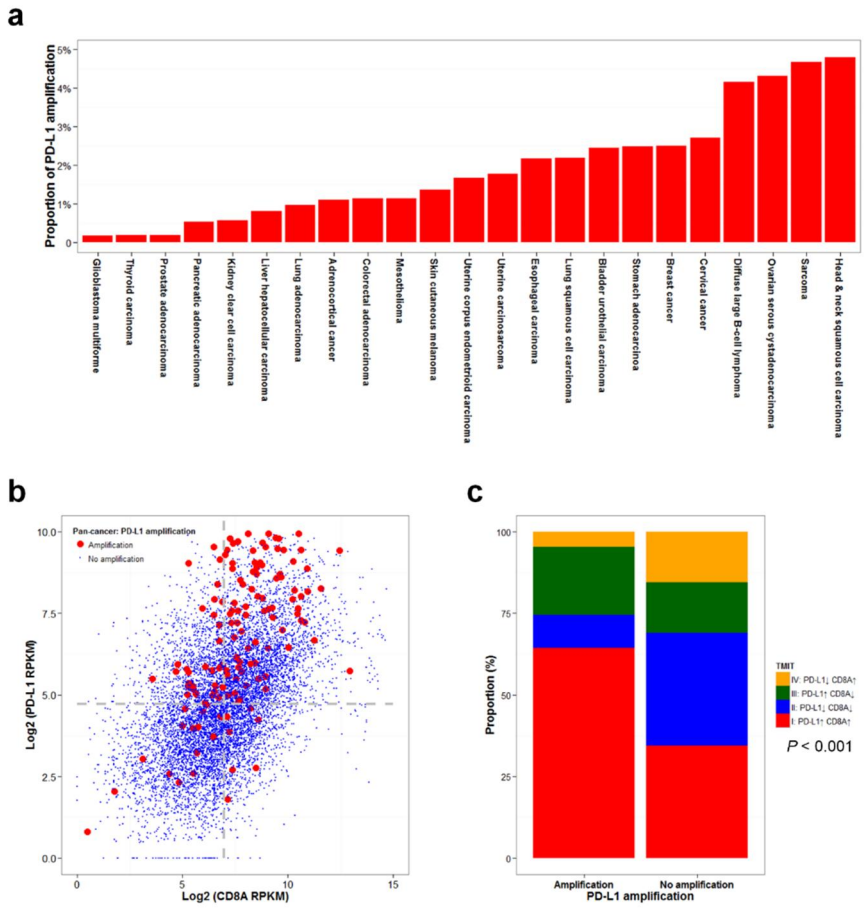

Figure 3. PD-L1 amplification is associated with tumor microenvironment immune type I. The frequency of PD-L1 amplification according to cancer type is shown (a). A scatter plot of log 2-transformed values of RPKM of PD-L1 and CD8A according to PD-L1 amplification is shown (b). The proportion of tumor microenvironment immune types according to PD-L1 amplification is compared (c). Abbreviation: RPKM, reads per kilobase of transcript per million reads mapped.

Oncogenic virus infection is associated with TMIT I

We then aimed to determine whether there is association between infection with an oncogenic virus and TMIT. Considerable proportions of EBV in STAD, HPV in CESC and HNSC, and hepatitis B virus (HBV) in liver hepatocellular carcinoma were detected (LIHC), which is consistent with previous reports (10, 49). A high proportion of tumors resulting from EBV infection and HPV infection were classified as TMIT I compared with those resulting from HBV infection or those with no virus (84% and 56.4% versus 18.5% and 35.9%, respectively, $P < 0.001$, **Figs. 1.4a** and **1.4b**). Interferon-gamma expression and cytolytic activity were significantly higher in TMIT I tumors (**Figs. 1.4c** and **1.4d**), and cytolytic activity was significantly higher in EBV-infected and HPV-infected tumors compared with HBV-infected and uninfected samples (**Fig. 1.4e**).

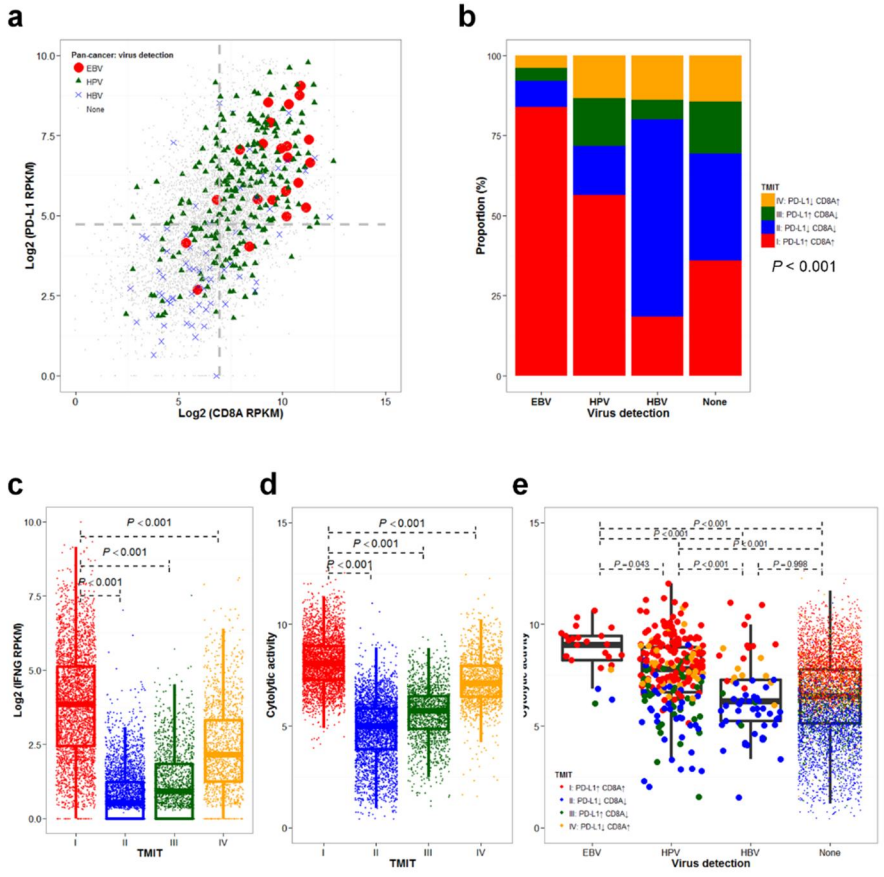

Figure 4. Oncogenic viral infection with increased cytolytic activity is associated with tumor microenvironment immune type I. A scatter plot of log 2-transformed values of RPKM of PD-L1 and CD8A according to virus detection and PD-L1 amplification is shown (a). The proportion of tumor microenvironment immune types (TMIT) according to PD-L1 amplification is compared (b). Box plots of log 2-transformed values of RPKM of interferon-gamma (IFNG, c) and cytolytic activity (d) according to TMITs are shown. A box plot of cytolytic activity according to virus detection and TMIT is shown (e).

Summary of clinicopathologic features correlated with TMIT I

Taken together, the multivariate analysis results showed that a high number of mutations, EBV infection, HPV infection, and *PD-L1* amplification were independently associated with TMIT I (**Table 2**). Biological impact of TMIT classification was summarized in **Fig. 5**.

Table 2. Logistic regression analysis for predicting tumor microenvironment immune type I according to clinicopathological characteristics

		Univariate		Multivariate	
		OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age	≥ Median	1.24 (1.14-1.35)	5.78×10^{-7}	1.11 (0.96-1.28)	0.147
Gender	Men (vs. Women)	0.97 (0.89-1.06)	0.493	not entered	not entered
Number of mutations	≥ Median	1.55 (1.40-1.72)	$< 2.00 \times 10^{-16}$	1.44 (1.25-1.67)	6.21×10^{-7}
Presence of virus	EBV	8.14 (3.07-28.0)	1.32×10^{-4}	6.29 (2.34-21.8)	9.01×10^{-4}
	HPV	2.31 (1.75-3.07)	5.05×10^{-9}	1.98 (1.47-2.67)	8.31×10^{-6}
	HBV	0.19 (0.07-0.42)	1.54×10^{-4}	0.18 (0.07-0.40)	1.08×10^{-4}
PD-L1 amplification	Yes	3.46 (2.50-4.87)	6.47×10^{-13}	3.45 (1.96-6.33)	3.10×10^{-5}

Abbreviation: EBV, Epstein-Barr virus; HBV, hepatitis B virus; HPV, human papillomavirus; OR, odds ratio.

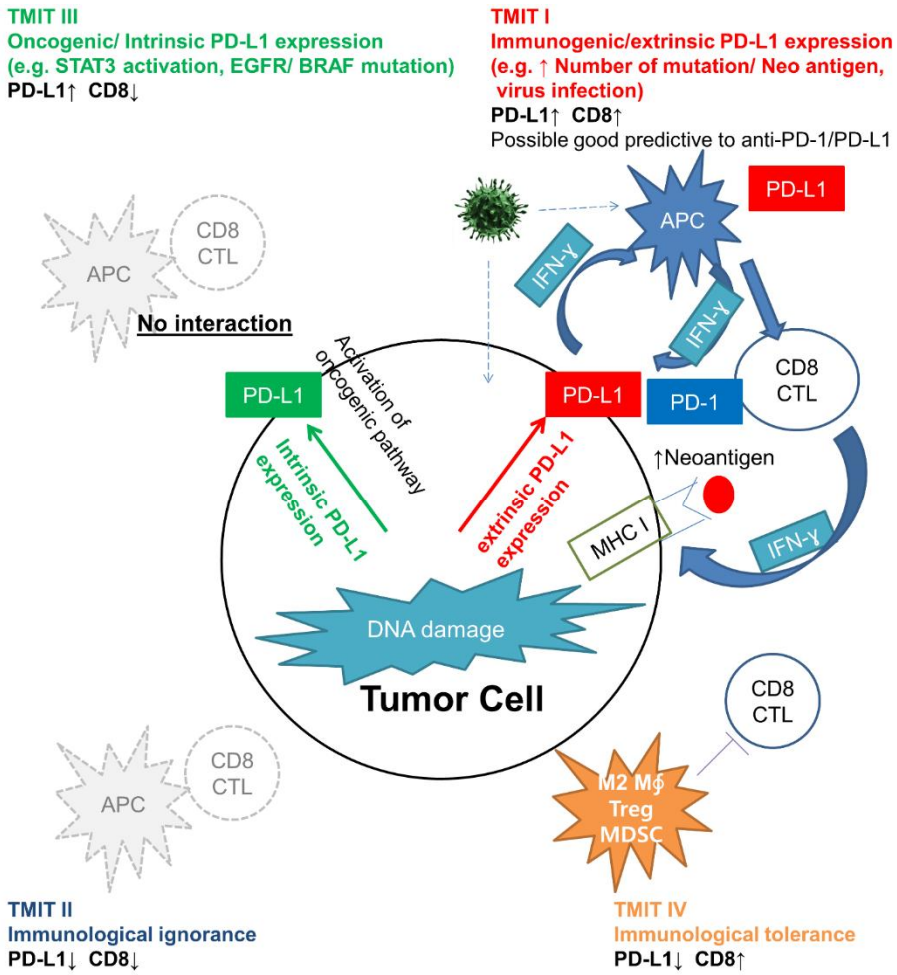

Figure 5. Summary of tumor microenvironment immune types.

Immune-suppressive signature in inflamed tumors

To understand immune-suppressive gene signature in inflamed tumors with highly recruited TIL, we divided TCGA tumor samples by different classification system. Using Cibersort algorism (36), infiltrated proportions of immune cells in TME can be predicted *in silico*. TIL-high TME is defined by higher CD8-positive T-cells infiltration score (CD8TS) predicted by Cibersort than their median value in all TCGA samples. As shown in **Fig 6**, CD8TS was linearly correlated with regulatory T-cells infiltration score (TregS) or M2 macrophage infiltration score (M2S), indicating that recruitment of TIL would consist of both immune-activators and immune-suppressors. Divided by 0.5 standard deviation higher than fitted value of TregS or M2S by CD8TS, we classified tumors with 3 groups each for TregS or M2, respectively: CD8TS-low, CD8TS-high/TregS-high, and CD8TS-high/TregS-medium-low (**Fig 6a**), or CD8TS-low, CD8TS-high/M2-high, and CD8TS-high/M2-medium-low (**Fig 6b**). This classification clearly demonstrated prognostic significance, as survival rate of CD8TS-low group was similar with that of CD8TS-high/TregS-high group ($P < 0.001$, **Fig 6c**). However, CD8TS-high/TregS-medium-low group had favorable survival compared to other 2 groups. Classification according to CD8TS and M2 had also similar results with CD8TS and TregS (**Fig 6d**). Therefore, for precise assessment of TME, not only CD8-positive T cells infiltration, but also regulatory T-cells and M2

macrophage infiltration should be considered since those proportion clearly affected tumor biology.

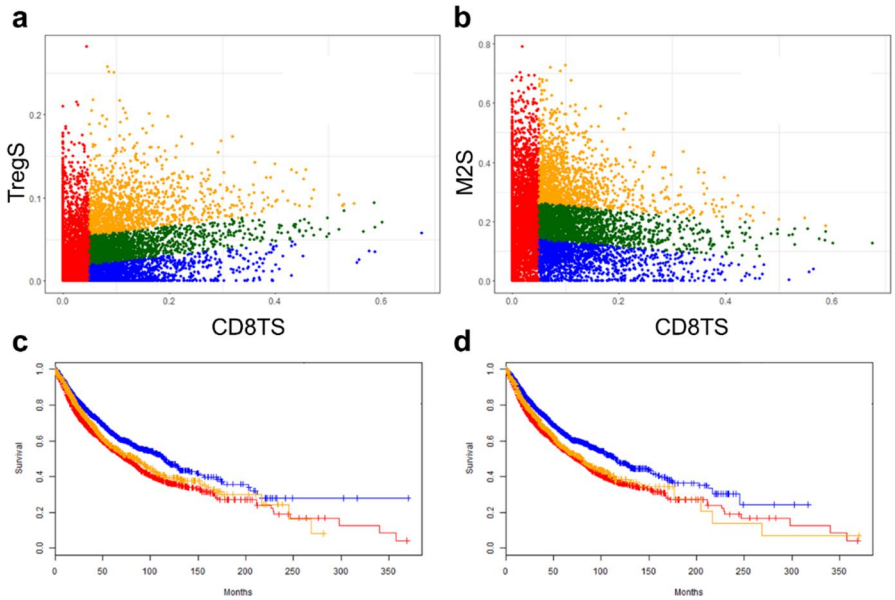

Figure 6. Biologic impact of immune-suppressive signature of regulatory T-cells and M2 macrophage. A scatter plot of CD8-positive T cells score (CD8TS, x-axis), and regulatory T-cells score (TregS, y-axis, a) and M2 macrophage (M2S, y-axis, b) is shown. Red dots indicates CD8TS-low group, and yellow dots indicates CD8TS-high/TregS-high (b), and CD8TS-high/M2S-high group (c). Green dots and blue dots indicate CD8TS-high/TregS (M2S)-medium and CD8TS-high/TregS (M2S)-low group, combining together as CD8TS-high/TregS (M2S)-medium-low group. Kaplan-Meier curves of overall survival according to CD8TS-low (red), CD8TS-high/TregS-high (yellow), and CD8TS-high/TregS-medium-low (blue) was shown (c). Kaplan-Meier curves of overall survival according to CD8TS-low (red), CD8TS-high/M2S-high (yellow), and CD8TS-high/M2S-medium-low (blue) was also shown (d).

EMT phenomenon is highly enriched in TMIT III

As previous studies shown that sarcoma and sarcomatoid lung carcinoma had high PD-L1 expression (25, 26), we hypothesized that EMT would be frequently observed in TMIT I or TMIT III cancers. To prove the proportion of EMT in TCGA, we defined that high vimentin mRNA expression and low E-cadherin mRNA expression would relate to EMT phenomenon. Among TCGA cancer types, thymoma (THYM), diffuse large B cell lymphoma (DLBC), acute myeloid leukemia (LAML), sarcoma (SARC), mesothelioma (MESO), glioblastoma (GBM), and germ cell tumor (TGCT) are classified as mesenchymal origin, and others are classified as epithelial origin. The ratio of vimentin and E-cadherin was increased in cancers with mesenchymal origin, compared to those with epithelial origin (**Fig 7a**). Interestingly, TMIT III had the most high level of vimenin / E-cadherin ratio among four TMITs (**Fig 7b**), indicating that sore PD-L1 expression would be closely related to EMT.

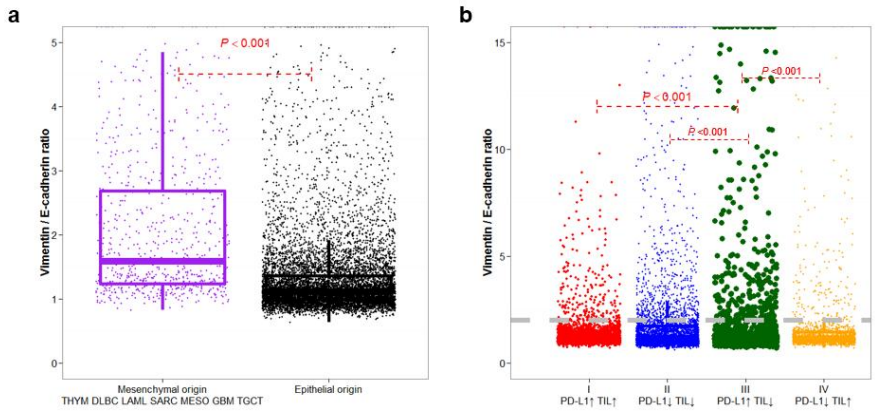

Figure 7. EMT is associated with TMIT III. Box plots of the ratio of vimentin and E-cadherin RPKM according to mesenchymal or epithelial origin (a) and according to TMIT (b).

PD-L1 expression correlated with EMT

Representative images of PD-L1-negative and -positive, p16-positive and high E-cadherin expressing tumor tissues are shown in Supplementary Figure 1. Of the 50 patients included in this study (training cohort), 32 and 18 patients, respectively, exhibited PD-L1-positive and –negative tumors. Fifteen patients were p16-positive and 17 were EMT-positive. Interestingly, 15 of the 32 PD-L1-positive tumors (46.9%) were EMT-positive, as assessed by low E-cadherin and high vimentin expression (**Fig. 8a**). PD-L1 positivity was significantly higher in patients with EMT-positive tumors (P = 0.013, **Fig. 8b**). Clinical features, including age, sex, smoking history, nd stage, did not differ according to patients' PD-L1 and EMT statuses (**Table 3**, left column). The proportion of patients showing oropharyngeal tumors was higher in PD-L1-positive/EMT-negative patients (PD-L1+/EMT-; Table 2.1). The E-cadherin H-score was lower for PD-L1+/p16- compared with PD-L1-/p16- patients (P = 0.559). The vimentin H-score was significantly higher in PD-L1+/p16- compared with PD-L1-/p16- patients (P = 0.014). However, this trend was not observed in p16+ patients (P = 0.245 and 0.371, respectively; **Fig. 8c-d**). Taken together, of the 32 PD-L1+ patients, 12 were p16-/EMT+ (37.5%), 9 were p16+/EMT- (28.1%), 8 were p16-/EMT- (25%), and 3 were p16+/EMT+ (9.4%; **Fig. 8e**). Although PD-L1 positivity was not significantly different according to p16 status (P = 0.199), PD-L1

positivity was significantly higher in p16-/EMT+ and p16+/EMT- compared with p16-/EMT- patients (P = 0.002 and 0.026, respectively). Univariate and multivariate logistic regression analyses showed that an oropharyngeal tumor origin and EMT status associated significantly with PD-L1 positivity (P = 0.014 and 0.010, respectively; **Table 4**).

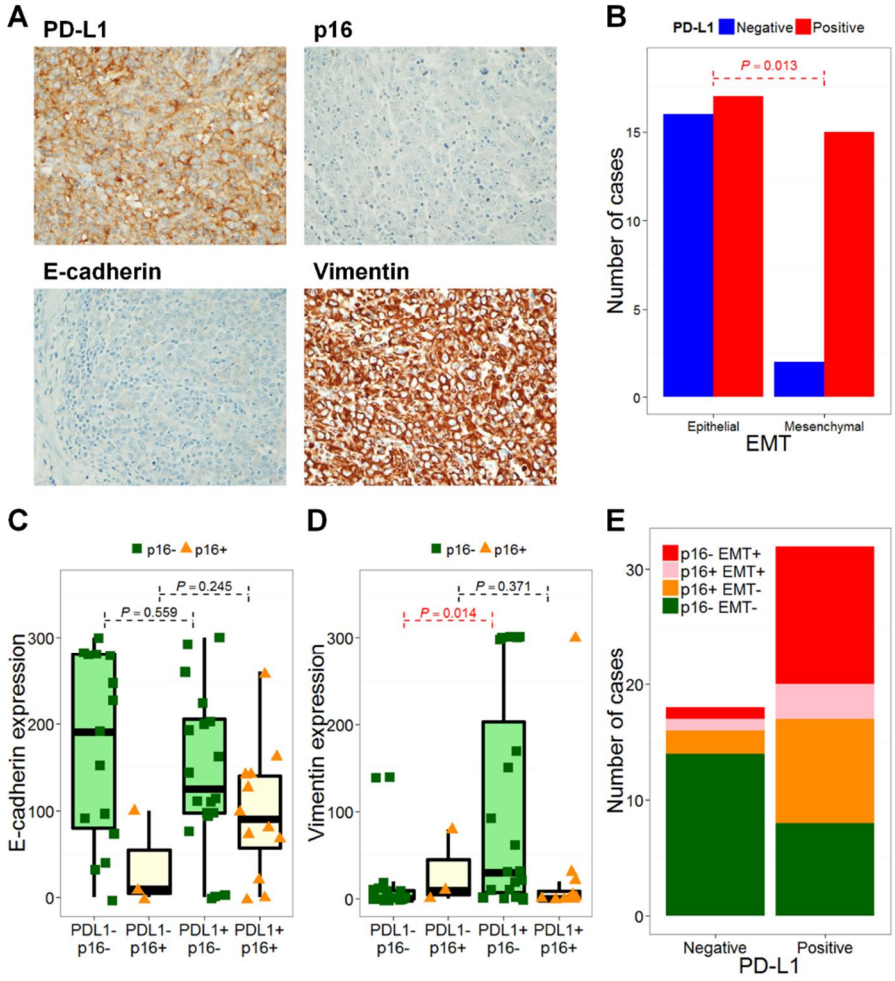

Figure 8. PD-L1 expression is associated with epithelial-mesenchymal transition. Representative images of immunohistochemical staining with anti-PD-L1 (x200; A, x400, B), anti-E-cadherin (x200, C), and anti-vimentin (x200, D) are shown. PD-L1 expression was positively correlated with vimentin and negatively correlated with E-cadherin. The number of PD-L1-negative (blue bars) and -positive (red bars) cases according to epithelial-mesenchymal transition (EMT) are shown. The P value from Fisher's exact test is annotated (red). H-scores for E-cadherin (C) and vimentin (D) are plotted according to PD-L1 and p16 statuses. The number of p16- EMT- (green), p16+ EMT- (orange), p16+ EMT+ (pink), and p16- EMT+ (red) cases according to PD-L1 status are shown (E).

Table 3. Patient characteristics according to PD-L1 and EMT status

		All N=50	PD-L1(-) N=18	PD-L1(+) EMT(-) N=17	PD-L1(+) EMT(+) N=15	*P* value
Age	Median years (range)	60 (16-78)	61 (44-78)	57 (16-75)	61 (26-76)	0.074
Sex	Men, N (%)	40 (80.0)	14 (77.8)	15 (88.2)	11 (73.3)	
	Women, N (%)	10 (20.0)	4 (22.2)	2 (11.8)	4 (26.7)	0.614
Smoking	Non-smoker, N (%)	30 (60.0)	10 (55.6)	10 (58.8)	10 (66.7)	
	Ex/Current-smoker, N (%)	20 (40.0)	8 (44.4)	7 (41.2)	5 (33.3)	0.877
ECOG	0, N (%)	12 (24.0)	2 (11.1)	6 (35.3)	4 (26.7)	
	1, N (%)	38 (76.0)	16 (88.9)	11 (64.7)	11 (73.3)	0.243
Location	Oropharynx, N (%)	16 (32.0)	2 (11.1)	10 (58.8)	4 (26.7)	
	Non-oropharynx, N (%)	34 (68.0)	16 (88.9)	7 (41.2)	11 (73.3)	**0.010**
p16	Negative, N (%)	35 (70.0)	15 (83.3)	8 (47.1)	12 (80.0)	
	Positive, N (%)	15 (30.0)	3 (16.7)	9 (52.9)	3 (20.0)	0.051
Pathology	P/D, N (%)	24 (48.0)	8 (44.4)	11 (64.7)	5 (33.3)	
	M/D, N (%)	12 (24.0)	6 (33.3)	3 (17.7)	3 (20.0)	
	W/D, N (%)	12 (24.0)	4 (22.2)	2 (11.8)	6 (40.0)	
	Non-keratinizing type, N (%)	2 (4.0)	0 (0.0)	1 (5.9)	1 (6.7)	0.331

Stage	I, N (%)	6 (12.0)	1 (5.6)	1 (5.9)	4 (26.7)	
	II, N (%)	2 (4.0)	0 (0.0)	1 (5.9)	1 (6.7)	
	III, N (%)	15 (30.0)	9 (50.0)	5 (29.4)	2 (13.3)	
	IVA, N (%)	27 (54.0)	8 (44.4)	10 (58.8)	8 (53.3)	0.253
Definitive Treatment	Concurrent chemoradiotherapy, N (%)	16 (32.0)	7 (38.9)	4 (23.5)	5 (33.3)	
	Surgery, N (%)	34 (68.0)	11 (61.1)	13 (76.5)	10 (66.7)	0.693
Overall survival	Median months (95% CI)	NR (43.7-NR)	50.1 (25.0-NR)	NR (NR-NR)	35.7 (30-NR)	**0.007**
	3-year survival rate	73.7%	74.3%	100%	42.8%	
	5-year survival rate	53.2%	41.3%	100%	21.4%	
Median follow-up	Median months (range)	72.4 (23.0-119.6)	84.5 (27.4-119.6)	50.3 (33.0-112.7)	48.2 (23-112.7)	0.623

Bold values indicate statistically significant correlations with *P* values less than 0.05.

Abbreviation: EMT, epithelial-mesenchymal transition; ECOG, Eastern Cooperative Oncology Group performance status; P/D, poorly-differentiated squamous cell carcinoma; M/D; moderate-differentiated squamous cell carcinoma; W/D, well-differentiated squamous cell carcinoma; CI, confidence interval; NR, not reached.

Table 4. Univariate and multivariate logistic analysis of factors affecting baseline PD-L1 expression

		Univariate		Multivariate	
		HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age	Continuous	0.97 (0.92-1.02)	0.275		
Sex	Female (vs. Male)	0.81 (0.19-3.35)	0.769		
Smoking	Yes (vs. No)	0.75 (0.23-2.42)	0.631		
ECOG	1 (vs. 0)	0.28 (0.05-1.43)	0.125		
Stage	Continuous	0.81 (0.44-1.49)	0.495		
Location	Oropharynx (vs. Non-oropharynx)	6.22 (1.22-31.7)	**0.028**	8.60 (1.54-48.1)	**0.014**
p16	Positive (vs. Negative)	3.00 (0.72-12.5)	0.132		
EMT	Positive (vs. Negative)	7.06 (1.39-35.9)	**0.018**	9.54 (1.72-52.9)	**0.010**

Bold values indicate statistically significant correlations with *P* values less than 0.05.

Abbreviation: ECOG, Eastern Cooperative Oncology Group performance status; HR, hazard ratio; CI, confidence interval; EMT, epithelial-mesenchymal transition.

Survival analysis according to PD-L1 and EMT statuses

Survival analysis according to PD-L1 expression in HNSCC has not been clearly defined. PD-L1 expression was not significantly associated with overall survival (OS, P = 0.137) or progression-free survival (PFS, P = 0.213). In regard to the training cohort, interestingly, PD-L1+/EMT+ patients showed significantly poorer OS and PFS rates compared to PD-L1+/EMT- patients (P < 0.001 and 0.005, respectively; **Fig. 9a-b**). The 3-year OS rate was 42.8% for PD-L1+/EMT+ patients, which differed markedly from 100% for PD-L1+/EMT- patients. Moreover, PD-L1+/EMT+ patients showed significantly poorer OS and PFS rates compared with PD-L1+/p16+/EMT- patients (P = 0.007 and 0.006, respectively).

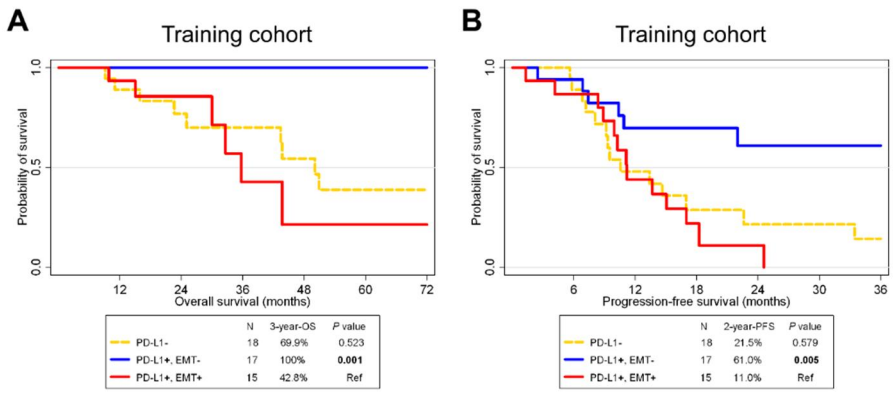

Figure 9. Survival analysis according to PD-L1 and epithelial-mesenchymal transition statuses in HNSCC patients. A Kaplan-Meier plot of overall survival (OS; A) and progression-free survival (PFS; B) according to PD-L1 and epithelial-mesenchymal transition (EMT) statuses. Abbreviations: HNSCC, head and neck squamous cell carcinoma; Ref, reference.

Correlation of PD-L1 with EMT in TCGA and CCLE

A 75 gene EMT expression signature was obtained from the GSE4824 dataset, which predicted mesenchymal features in lung cancer cell lines, including squamous cell carcinoma (40). EMT predictions based on the 75 gene signature were conducted for HNSCC samples and cell lines from TCGA and CCLE, respectively (28, 39). EMT predictions resulted in 119 out of 564 HNSCC samples (21.1%) from TCGA, and 4 out of 32 cell lines (12.5%) from CCLE having mesenchymal features (**Fig. 10a, c**). Samples with a high probability of mesenchymal features exhibited a high expression of mesenchymal signatures, such as *ZEB1* and *VIM* that encode vimentin, and low expression of epithelial signatures, such as *MUC1* and CDH1 that encode E-cadherin. Interestingly, in both TCGA and CCLE, PD-L1 expression was significantly higher in mesenchymal features compared with epithelial features ($P < 0.001$; **Fig. 10b, d**).

OS and PFS rates for PD-L1+/EMT+ patients were significantly worse compared to those for PD-L1- patients from TCGA ($P = 0.017$ and 0.009, respectively; **Fig. 11**). PFS rates were significantly different between PD-L1+/EMT+ and PD-L1+/EMT- patients ($P = 0.040$). However, OS rates were not significantly different, although the 3-year OS rate differed between PD-L1+/EMT+ (40.2%) and PD-L1+/EMT- (61.8%; $P = 0.109$).

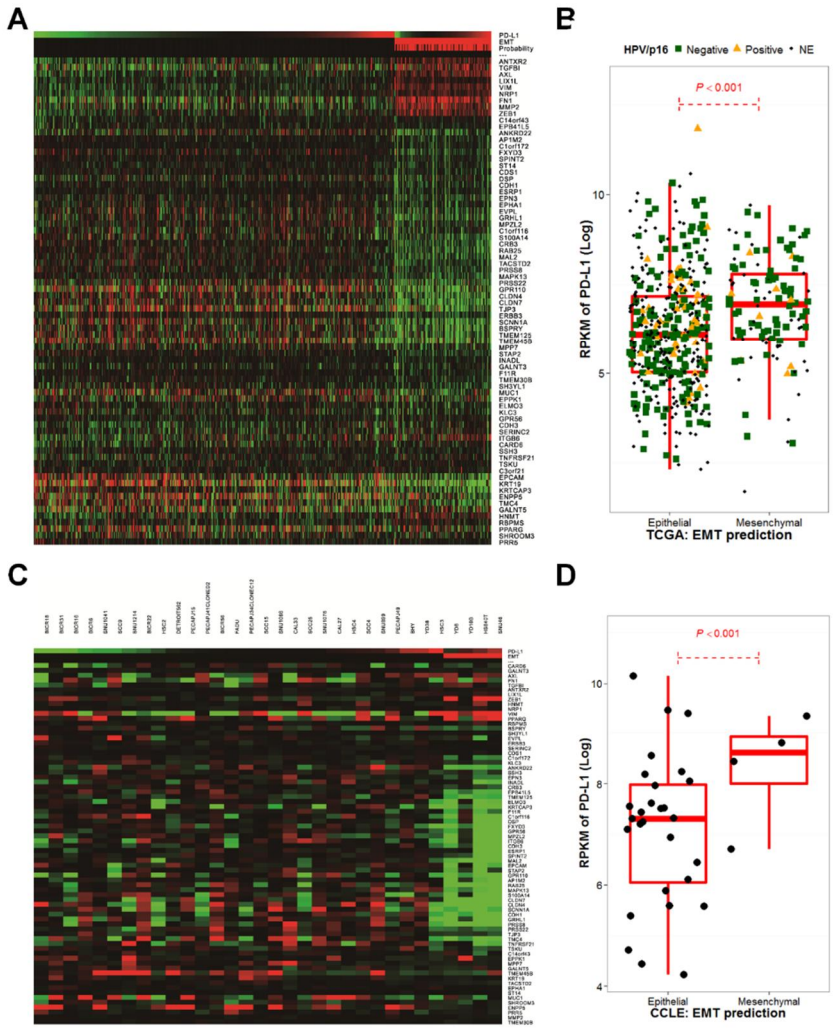

Figure 10. The epithelial-mesenchymal transition gene expression signature correlates with PD-L1 expression in The Cancer Genome Atlas and the Cancer Cell Line Encyclopedia. Significant 75-gene expression signatures, referred from GSE4824 and Bayesian probability to predict EMT changes that favor mesenchymal features in The Cancer Genome Atlas (TCGA) (A) and the Cancer Cell Line Encyclopedia (CCLE) cohorts (C) as well as PD-L1 expression, are shown as a heatmap. PD-L1 expression was calculated by the log 2 value of its reads per kilobase of transcript per million mapped reads (RPKM) and was compared according to EMT predictions in TCGA (B) and CCLE cohorts

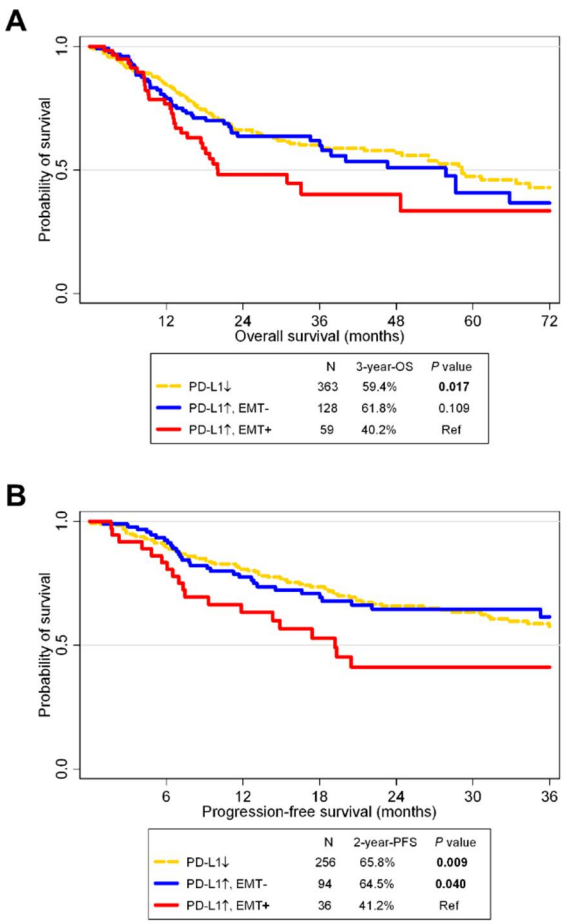

Figure 11. Survival analysis according to PD-L1 and epithelial-mesenchymal statuses in The Cancer Genome Atlas cohort. A Kaplan-Meier plot of overall survival (OS; A) and progression-free survival (PFS; B) according to PD-L1 expression and epithelial-mesenchymal transition (EMT) status. Abbreviation: Ref,

PD-L1 expression changes in HNSCC patient samples

Of the 35 HNSCC patients included to analyze whether PD-L1 expression would change by treatment, baseline tumor tissue of 22 patients (62.9%) had PD-L1-positive tumors, while the other 13 patients (37.1%) had PD-L1-negative tumors (**Table 5**). Clinical features including tumor location of the oropharynx and p16-status did not significantly differ according to baseline PD-L1 status, although baseline PD-L1-positive tumors showed a relatively high proportion of oropharyngeal tumors (7 of 22, 31.8% in PD-L1-positive vs. 1 of 13, 7.7% in PD-L1-negative) and p16-positive tumors (6 of 22, 27.3% in PD-L1-positive vs. 1 of 13, 7.7% in PD-L1-negative).

During the treatment course, PD-L1 changes in various ways. Of the 13 patients who were PD-L1-negative prior to treatment, 9 cases (69.2%) showed up-regulated PD-L1 expression after treatment (P = 0.003). In contrast, 4 of 22 patients (18.2%) who were initially PD-L1-positive tumor showed decreased PD-L1 expression (P = 0.072, **Fig. 12a-b**). Interestingly, most patients who underwent cisplatin treatment as induction chemotherapy or concurrent chemoradiotherapy (CCRT) showed up-regulated PD-L1 expression in their post-treatment tumor tissue (P = 0.037, **Fig. 12c** and **Table 6**).

Interestingly, E-cadherin expression was decreased (P = 0.006) and vimentin expression was generally increased (P = 0.393) in post-treatment tumor tissue compared to in matched baseline tissues (**Fig.**

13). However, we found no association between PD-L1 changes and morphological EMT occurrence such as sarcomatoid like change. Moreover, intervals of biopsies between baseline and after treatment was not associated with PD-L1 changes.

Table 5. Patient characteristics according to PD-L1 status before and after treatment

PD-L1	**Before treatment**	**Negative**	**Negative**	**Positive**	**Positive**	**All**	***P***	***P***
	After treatment	**Negative**	**Positive**	**Negative**	**Positive**			
		N = 4	**N = 9**	**N = 4**	**N = 18**	**N = 35**	***Baseline PD-L1****	***4 groups***
Age	Median years (range)	68 (52-78)	63 (51-70)	42 (26-76)	62 (16-75)	63 (16-78)	0.260	0.443
Sex	Men, N (%)	3 (75.0)	8 (88.9)	4 (100.0)	16 (88.9)	31 (88.6)		
	Women, N (%)	1 (25.0)	1 (11.1)	0 (0)	2 (11.1)	4 (11.4)	0.478	0.860
Smoking	Non-smoker, N (%)	2 (50.0)	5 (55.6)	3 (75.0)	14 (77.8)	24 (68.6)		
	Ex/Current-smoker, N (%)	2 (50.0)	4 (44.4)	1 (25.0)	4 (22.2)	11 (31.4)	0.144	0.528
ECOG	0, N (%)	0 (0)	1 (11.1)	1 (25.0)	8 (44.4)	10 (28.6)		
	1, N (%)	4 (100)	8 (88.9)	3 (75.0)	10 (55.6)	25 (71.4)	0.055	0.078
Location	Oropharynx, N (%)	0 (0)	1 (11.1)	2 (50.0)	5 (27.8)	8 (22.9)		
	Non-oropharynx**, N (%)	4 (100)	8 (88.9)	2 (50.0)	13 (72.2)	27 (77.1)	0.108	0.346

p16	Positive, N (%)	0 (0)	1 (11.1)	3 (75.0)	3 (16.7)	7 (20.0)		
	Negative, N (%)	4 (100)	8 (88.9)	1 (25.0)	15 (83.3)	28 (80.0)	0.170	0.058
Pathology	SqCC P/D, N (%)	2 (50.0)	3 (33.3)	1 (25.0)	6 (33.3)	12 (34.3)		
	SqCC M/D, N (%)	1 (25.0)	5 (55.6)	2 (50.0)	2 (11.1)	10 (28.6)		
	SqCC W/D, N (%)	1 (25.0)	1 (11.1)	1 (25.0)	8 (44.4)	11 (31.4)		
	Non-keratinizing type, N (%)	0 (0)	0 (0)	0 (0)	2 (11.1)	2 (5.7)	0.189	0.375
Stage	I, N (%)	1 (15.0)	0 (0)	1 (25.0)	3 (16.7)	5 (14.3)		
	II, N (%)	0 (0)	0 (0)	0 (0)	1 (5.6)	1 (2.9)		
	III, N (%)	3 (75.0)	2 (22.2)	0 (0)	2 (11.1)	7 (20.0)		
	IVA, N (%)	0 (0)	7 (77.8)	3 (75.0)	12 (66.7)	22 (62.9)	0.187	0.254
Definitive treatment	Concurrent chemoradiotherapy, N (%)	3 (75.0)	4 (44.4)	1 (25.0)	6 (33.3)	14 (40.0)		
	Surgery, N (%)	1 (25.0)	5 (55.6)	3 (75.0)	12 (66.7)	21 (60.0)	0.177	0.507
Cisplatin treatment	No (never), N (%)	3 (75.0)	4 (44.4)	3 (75.0)	8 (44.4)	18 (51.4)		

	Yes (ever), N (%)	1 (25.0)	5 (55.6)	1 (25.0)	10 (55.6)	17 (48.6)	0.552	0.212
Interval between harvesting tissues	Median months (range)	12.4 (1.0-44.5)	14.9 (1.5-33.4)	9.4 (3.3-11.6)	8.7 (1.8-39.9)	11.6 (1.0-44.5)	0.246	0.603
Overall survival	Median months (95% CI)	25.0 (15.9-NR)	50.1 (22.6-NR)	NR (35.7-NR)	43.7 (30-NR)	50.1 (32.6-NR)	0.451	0.859
	3-year survival rate	37.5%	72.9%	66.7%	61.5%	63.0%		
	5-year survival rate	37.5%	18.2%	0%	49.2%	38.6%		
Median follow-up	Median months (range)	75.1 (46.3-88.2)	62.4 (27.4-119.6)	73.7 (23-111.7)	45.1 (29-104.5)	62.4 (23-119.6)	0.306	0.641

* *P* value of comparison of PD-L1-negative before treatment (1st and 2nd columns) and PD-L1-positive before treatment (3rd and 4th columns).

** Non-oropharynx included hypopharynx, larynx, nasal cavity, paranasal sinus, and oral cavity, which were not significant according to PD-L1 positivity.

Abbreviation: ECOG, Eastern Cooperative Oncology Group performance status; SqCC, Squamous cell carcinoma; P/D, poorly-

differentiated squamous cell carcinoma; M/D; moderate-differentiated squamous cell carcinoma; W/D, well-differentiated squamous cell carcinoma; CI, confidence interval; NR, not reached.

Table 6. Summary of PD-L1-negative HNSCC patient treated with cisplatin

HNSCC type	Sex/Age	p16 status	Brief history	Baseline tissue	Baseline PD-L1	Post-treatment tissue	Post PD-L1
hypopharynx	F/66	(-)	Induction chemotherapy → definitive CCRT → salvage operation	Before induction chemotherapy, primary tumor	(-)	On salvage operation, recurred tumor	(-)
Hypopharynx	M/63	(-)	Induction chemotherapy → definitive CCRT → salvage operation	Before induction chemotherapy. primary tumor	(-)	On salvage operation, recurred tumor	(1+)
Pyriform sinus	M/62	(-)	Induction chemotherapy → definitive operation → post-RT → salvage operation	Before induction chemotherapy. primary tumor	(-)	On salvage operation, recurred tumor	(1+)
Nasal cavity	M/60	(-)	Definitive CCRT → salvage operation	Before CCRT, primary tumor	(-)	On salvage operation, recurred tumor	(1+)
Larynx	M/66	(-)	Definitive operation → post-CCRT → salvage operation	On definitive operation, primary tumor	(-)	On salvage operation, recurred tumor	(1+)
Hypopharynx	M/69	(-)	Induction chemotherapy → definitive CCRT → salvage operation	Before induction chemotherapy. primary tumor	(-)	On salvage operation, recurred tumor	(1+)

Abbreviation: HNSCC, head and neck squamous cell carcinoma; CCRT, concurrent chemoradiotherapy; RT, radiotherapy.

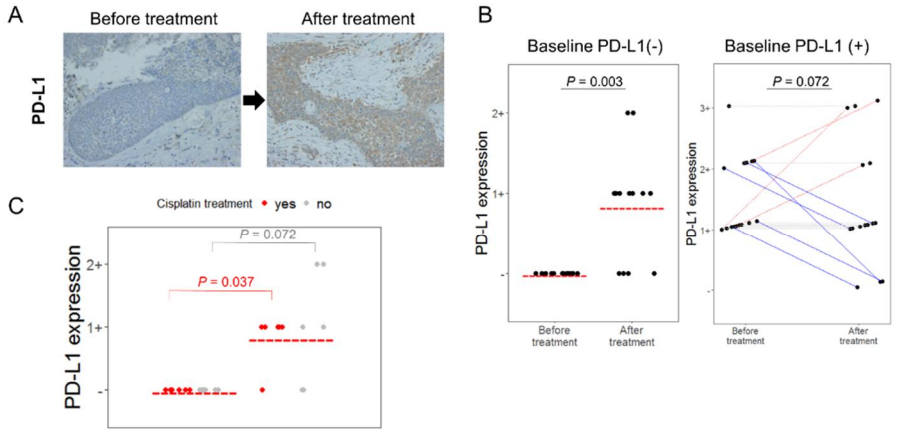

Figure 12. PD-L1 changed after treatment in HNSCC patient samples.

A, Representative PD-L1 immunochemical staining (x400). PD-L1 expression was increased after treatment. B, Changes in PD-L1 expressions before treatment (baseline) and after treatment in 35 HNSCC patients are shown. Each dot represents PD-L1 expression positivity, and red bar represented indicates PD-L1 expression positivity in each group. For the right side graph, each line connects the same individual. Red dash lines represent PD-L1 positivity is increased, while blue solid lines represents PD-L1 positivity is decreased. C, Among baseline PD-L1-negative patients, PD-L1 positivity measured by immunohistochemistry was more significantly up-regulated in cisplatin-treated patients (red dot) compared to those who were not exposed to cisplatin (grey dot). P values were noted for comparison of before treatment (baseline) PD-L1 positivity and after treatment PD-L1 positivity in each group of cisplatin-treated patients (red) and cisplatin-naïve patients (grey).

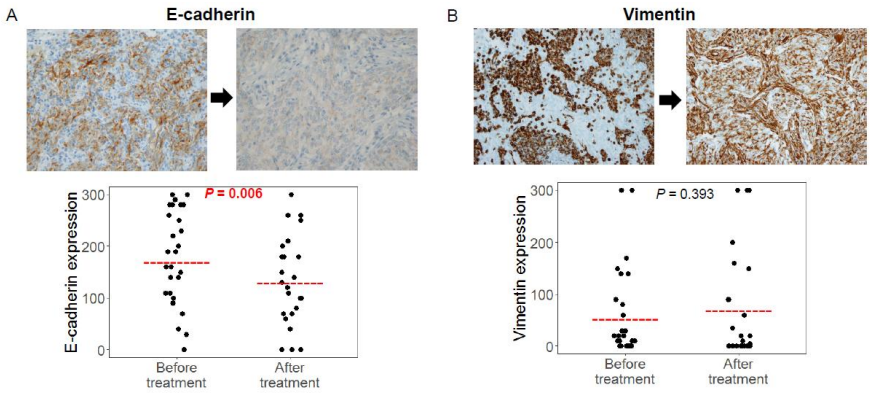

Figure 13. Changes of EMT markers before and after treatment.

Representative cases of e-cadherin down-regulation (A, x400, top) and vimentin up-regulation (B, x400, top) after treatment in HNSCC patient samples. Each dot represented E-cadherin (A) or vimentin (B) expression positivity, and red bar represented mean e-

Increased PD-L1 expression in response to cisplatin treatment *in vitro*

To confirm that PD-L1 expression changes following cisplatin treatment *in vitro*, HNSCC cell lines were treated with cisplatin for 24 h and PD-L1 expression after cisplatin treatment was compared with baseline expression. In all HNSCC cell lines analyzed by flow cytometry, PD-L1 expression was up-regulated after cisplatin treatment (**Fig. 14**).

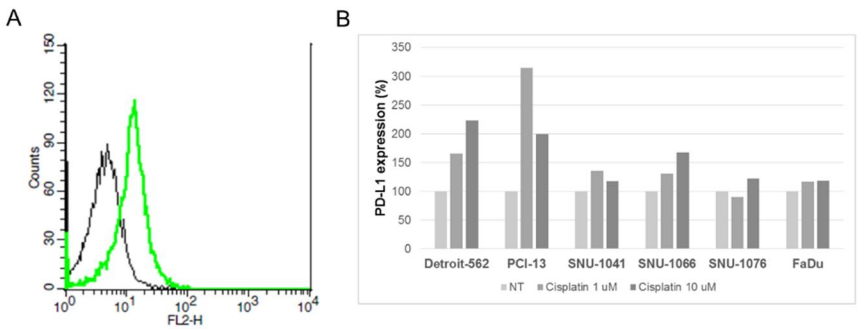

Figure 14. Increased PD-L1 expression by cisplatin treatment in HNSCC cells.
A, Cisplatin 1 microM (green line) or normal saline (black line) was used to treat SNU-1041 cells for 24 h, and then flow cytometry analysis was performed using anti-PD-L1 antibody. PD-L1 expression was measured as the geometrical mean of fluorescence in gated cells. B, Bar graph showing mean PD-L1 expressions according to cisplatin treatment in HNSCC cells. Each bar represents the percent change of PD-L1 expression compared to no treatment (NT) in each HNSCC cell line.

Increased PD-L1 expression accompanied activation of MEK pathway

To determine which signaling pathways are related to PD-L1 up-regulation by cisplatin, western blot analysis of phospho-MEK and phospho-STAT3 was performed and PD-L1 expression was determined. PD-L1 expression increased in all HNSCC cells according to western blot analysis. Interestingly, cisplatin treatment increased the ratio of phospho-MEK/total-MEK in a dose-dependent manner in PCI-13 and SNU-1066 cells. However, the ratio of phosphor-STAT3/total-STAT3 did not increase or decrease in all HNSCC cells according to cisplatin treatment (**Fig. 15**). It has been previously reported that EMT is associated with PD-L1 upregulation in gefitinib-resistance in NSCLC (47); thus, we evaluated the expression of EMT markers such as E-cadherin and vimentin in cisplatin-treated HNSCC cells. However, EMT expression did not change after cisplatin treatment in HNSCC cells (**Fig 16**).

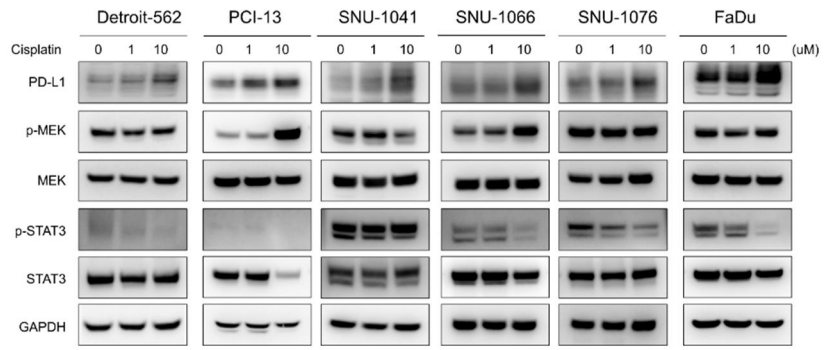

Figure 15. Increased PD-L1 expression by cisplatin accompanied MEK pathway activation in HNSCC cells.
Expression of PD-L1, phosphor-MEK (p-MEK), total MEK, phosphor-STAT3 (p-STAT3), and total STAT3 in HNSCC cells was measured by western blotting. Cisplatin treatment increased PD-L1 and p-MEK expression.

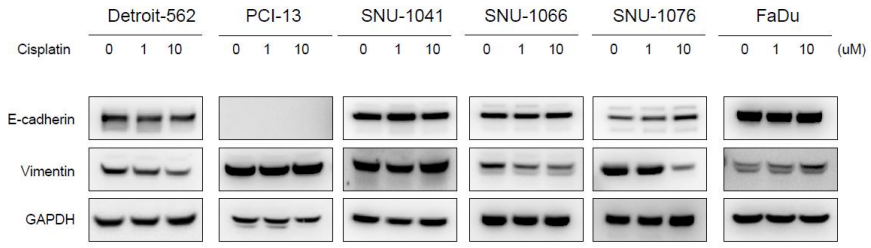

Figure 16. Association of PD-L1 expression and epithelial-mesenchymal transitions (EMT) markers in head and neck squamous cancer cells.
E-cadherin and vimentin expressions measured by western blot were compared according to cisplatin treatment in head and neck squamous cancer cells.

Discussion

Using a large-scale TCGA pan-cancer dataset, we classified all types of cancer into one of four TMITs based on their PD-L1 and CD8A mRNA expression levels assessed by RNA sequencing. TMIT varies across cancer type, and this result is likely accompanied by the prevalence of somatic mutations across the cancer types. TMIT I, defined by high PD-L1 and high CD8A expression, is associated with a high mutational burden/neoantigen, high MSI, *PD-L1* amplification, and infection with an oncogenic virus such as EBV and HPV. Our results suggest that PD-L1 positivity should be comprehensively interpreted with TME and that classification into four TMITs may be an appropriate model through which to tailor cancer immunotherapeutic modules (**Fig. 5**).

Consistent with previous observations, the proportion of TMIT I tumors, which reflects the adaptive immune response (15), was higher in lung cancer and melanoma, as well as head and neck cancer, stomach cancer, and bladder cancer, which are expected to have a favorable clinical outcome in response to anti-PD-1/PD-L1 therapies (5, 50, 51).

The assessment of the tumor mutational burden has indicated that the mutational burden varies between studies, and the prediction of neoantigens by calculating the interaction between a specific mutation and HLA genotype would be theoretically appropriate (10, 12, 52).

However, the number of non-synonymous mutations is also clearly correlated with the mutational burden as well as the clinical outcome of anti-PD-1 antibody treatment (8, 13). In the current study, the number of predicted neoantigens, as referenced by Rooney *et al*. (10), was found to be significantly correlated with the total number of somatic mutations assessed by whole-exome sequencing. Considering that the prediction of neoantigens is only performed *in silico (53, 54)*, calculating the number of somatic mutations would be a more convenient method for assessing the mutational burden because the total number of mutations is also significantly correlated with the number of non-synonymous mutations (data not shown).

Consistent with a previous report that showed that oncogenic virus infection increases the cytolytic activity of a tumor (10), a high proportion of tumors associated with EBV and HPV are TMIT I. Viral infection generates various viral antigens inside the tumor, and this phenomenon increases the immunogenicity of the tumor by activating the interferon-gamma (IFNG) pathway (55), as supported by recent findings of the association of PD-L1 with EBV-associated malignancies (20, 21). Moreover, *PD-L1* amplification is also clearly correlated with TMIT I. Previous results have clearly shown that Hodgkin's lymphoma presents variable copy number gains of chromosome 9p24.1, a genomic region that includes *PD-L1*, *PDCD1LG2* (encoding PD-L2, another ligand of PD-1), and *JAK2,* which activates the JAK/STAT/IFNG

pathway (18, 56). This finding is consistent with our analysis, which showed a frequent co-amplification of *PD-L1* and *JAK2* in the TCGA pan-cancer data (data not shown). These findings are also clearly supported by a recent finding of anti-PD-1 antibody in Hodgkin's lymphoma, which is known to be associated with both EBV and *PD-L1* amplification (4).

In this study, the proportion of TMIT I tumors was found to be correlated with the number of mutations across cancer type and viral infection status, but considerable discrepancies also exist. Kidney clear cell carcinoma (KIRC) is known as a highly immunogenic tumor type (57), and the results from this study show a high proportion of TMIT I. However, this tumor has a relatively low mutational burden. Interestingly, *VHL* and *PBRM1* mutations, which are frequently shown in KIRC (58), were clearly observed to be correlated with TMIT I. Moreover, hepatocellular carcinoma has a low proportion of TMIT I tumors, although it has a relatively high mutational burden and an association with HBV. These findings are likely the result of the unique anatomical characteristics of immune privilege, such as in glioblastoma (59, 60).

In the current study, we used RNAseq data from a mixture of cancer cells and surrounding tissues, including TILs, although this contamination of surrounding stromal tissues was not intentional. Interestingly, as interpreted by Rooney *et al.* (10), this limitation would

actually be advantageous to assessing the infiltration of surrounding tissues, such as the TIL of CD8+ CTLs. In this regard, the assessment of CD8A gene expression in a mixture of cancer and stromal cells would be more practical in the clinical setting than the immunohistochemical assessment of CD8+ CTL recruitment or activity, which would be difficult to judge uniformly across different tumor types. Nevertheless, the clinical validation of this approach is definitely warranted.

Another interesting finding of the current study is that immune-suppressive signature such as regulatory T-cells and M2 macrophage would be combined with immune-activator such as CD8-positive T-cells. It has been suggested that precise balance of immune-activator and immune-suppressor would result in overall anti-tumor effect of immune system, but this hypothesis has not been clearly defined. Using large-scale public database, TCGA, and smart algorism to assess the proportion of immune cells, Cibersort (36), we clearly proved that infiltration of CD8-positive T-cells is positively correlated with that of regulatory T-cells or M2 macrophage, and the precise ratio of immune activator and immune-suppressor would largely affect tumor biology, since their prognosis was clearly different according to the ratio. This approach would be far meaningful when we assess which TME would be favorable to immune checkpoint inhibitor, since highly-infiltrated regulatory T-cells or M2 macrophage would interfere anti-tumor effect

of the agent.

In the current study, EMT is clearly associated with TMIT III, which had high PD-L1 expression without TIL infiltration. In HNSCC, PD-L1 expression significantly correlated with EMT rather than HPV/p16 status. In addition to an association between PD-L1 and EMT in a rodent model (27) and a strong PD-L1 association with cancers of mesenchymal origin (25, 26), this is the first report to show a correlation between PD-L1 and EMT in cancers that originate from epithelial tissue. This correlation was significant both in clinical samples from a cohort of Korean patients and in an independent cohort that primarily consisted of Western patients in TCGA (28) and a comprehensive cancer cell line database, CCLE (39).

A statistical significance regarding PD-L1 associating with HPV/p16 status was not observed in clinical or TCGA samples. However, HPV/p16 could also partially contribute to PD-L1 expression given that oropharyngeal tumors, with their high prevalence of HPV infection, also correlated with PD-L1 positivity. HPV-positive HNSCC and other virus-associated cancers, such as HPV-positive uterine cervical cancer and EBV-positive gastric cancer, show increased immunogenic features such as an abundance of TIL or CD8-positive cytotoxic T cell signatures (10, 61). Because the correlation between PD-L1 and EMT is increased in HPV/p16-negative patients, HPV/p16-positivity and EMT features could contribute to PD-L1 expression in a

mutually exclusive manner.

Overall prognosis according to PD-L1 status is controversial due to various factors. PD-L1 expression is considered a poor prognostic factor in cancers such as non-small cell lung cancer, renal cell carcinoma, and melanoma (62-64). In contrast, PD-L1 status is a good prognostic factor in colorectal cancer (65). However, in regard to HNSCC, the lack of information in the literature concerning its prognosis according to PD-L1 expression could be due to a bias that PD-L1 is not important in this disease. In the current study, PD-L1 expression did not solely affect significant survival differences. However, the inclusion of factors associated with PD-L1 up-regulation, such as HPV/p16 positivity and EMT features, clearly defined distinct patient groups by survival differences. In PD-L1-positive patients, p16 positivity conferred a much better prognosis than EMT positivity given that viral associations induce immunogenic features in tumors, which confer a good prognosis in HNSCC (66, 67). Cancers showing EMT are associated with early recurrence and aggressive metastases (68); therefore, EMT-associated PD-L1-positive cancers have a poor prognosis. This prognostic trend was validated in a TCGA cohort that included diverse ethnicities.

Examining the underlying mechanism of PD-L1 up-regulation is challenging. Various reported have shown that the MEK pathway is activated according to anti-epidermal growth factor receptor treatment

resistance in NSCLC, and inhibition of the MEK pathway attenuated PD-L1 up-regulation (47, 69). Therefore, a combination of an MEK inhibitor and PD-L1 inhibitor may be synergistic (70), which has been evaluated in clinical trials. In the current study, we showed that MEK pathway activation was accompanied by PD-L1 up-regulation in cisplatin-treated HNSCC cells, indicating that MEK regulation is a crucial step in modulating PD-L1 expression in cancer.

Clinically, it is difficult to perform repeated biopsy considering bleeding risks. Therefore, assessing the immunologic profile of cancer, particularly PD-L1 status, is typically used to analyze tissues that are usually harvested by surgery or biopsy at initial diagnosis. However, as host-tumor immunologic status continuously changes, PD-L1 expression is also altered during treatment. We found that PD-L1 expression was altered in 37.1% of tumor samples after treatment. This trend was particularly strong in baseline PD-L1-negative patients (69.2%) and cisplatin-treated patients (83.3%). Previous clinical trials consistently showed that PD-L1 expression is an important biomarker for predicting anti-PD-1/PD-L1 treatment in various cancer types including HNSCC (71, 72); thus, PD-L1 assessment should be performed in recent biopsy samples in order to precisely determine whether anti-PD-1/PD-L1 inhibitors should be used. Well-designed proof-of-concept clinical trials to confirm whether recent biopsy samples should be used for accurate biomarker analysis are needed.

Although the current study has some limitations, including that the cutoff values of PD-L1 and CD8A need clinical validation, it is nevertheless valuable because we systematically assessed the TMIT of most cancer types by using data from TCGA project. Future investigations and clinical validations regarding the use of this approach for the assessment of immuno-genomic features across cancer types are warranted. In HNSCC studies, retrospective design and relatively small number of samples with a heterogeneous clinical status used in the current study could have biased results. However, significant statistical results were obtained for the independent validation cohort, TCGA and the CCLE databases. Although nearly a half of the patients received cisplatin chemotherapy, the sequence of treatment modalities and definitive treatment were very heterogeneous. However, the main finding of the study was that PD-L1 changes according to treatment, which is important because it is difficult to obtain paired biopsy samples because of technical and ethical issues. In-depth *in vitro* analysis to clearly define the association between the MEK pathway and PD-L1 up-regulation is needed.

In conclusion, from the pan-cancer immuno-genomic perspective, the classification of tumors into four TMITs based on PD-L1 status and CD8A+ TIL is an appropriate approach for cancer immunotherapy. TMIT I (PD-L1+/CD8A+) was found to be associated

with high PD-L1 expression, high mutational burden/neoantigen, high MSI, *PD-L1* amplification and the presence of an oncogenic virus. Next, PD-L1 expression is associated with EMT, an independent up-stream pathway distinct from HPV/p16 association. EMT-associated PD-L1 expression confers a significantly poorer prognosis compared to PD-L1 expression not associated with EMT. Moreover, PD-L1 expression in HNSCC is altered during the treatment phase, particularly following cisplatin-containing chemotherapy in baseline PD-L1-negative patients. Clinical investigations using anti-PD-1/PD-L1 inhibitors focused on TMIT classification, EMT-associated or chemotherapy-associated PD-L1 up-regulation are warranted.

References

1. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. The New England journal of medicine. 2010;363(8):711-23.
2. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. The New England journal of medicine. 2012;366(26):2443-54.
3. Hamid O, Robert C, Daud A, Hodi FS, Hwu WJ, Kefford R, et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. The New England journal of medicine. 2013;369(2):134-44.
4. Ansell SM, Lesokhin AM, Borrello I, Halwani A, Scott EC, Gutierrez M, et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. The New England journal of medicine. 2015;372(4):311-9.
5. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. The New England journal of medicine. 2015;372(21):2018-28.
6. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature. 2014;515(7528):563-7.
7. Taube JM, Klein A, Brahmer JR, Xu H, Pan X, Kim JH, et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. Clinical cancer research : an official journal of the American Association for Cancer Research. 2014;20(19):5064-74.
8. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-21.
9. Rutledge WC, Kong J, Gao J, Gutman DA, Cooper LA, Appin C, et al. Tumor-infiltrating lymphocytes in glioblastoma are associated with

specific genomic alterations and related to transcriptional class. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013;19(18):4951-60.

10. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015;160(1-2):48-61.

11. Brown SD, Warren RL, Gibb EA, Martin SD, Spinelli JJ, Nelson BH, et al. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. Genome research. 2014;24(5):743-50.

12. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science. 2015;348(6230):69-74.

13. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015;348(6230):124-8.

14. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. The New England journal of medicine. 2015;372(26):2509-20.

15. Teng MW, Ngiow SF, Ribas A, Smyth MJ. Classifying Cancers Based on T-cell Infiltration and PD-L1. Cancer research. 2015;75(11):2139-45.

16. Taube JM, Anders RA, Young GD, Xu H, Sharma R, McMiller TL, et al. Colocalization of inflammatory response with B7-h1 expression in human melanocytic lesions supports an adaptive resistance mechanism of immune escape. Science translational medicine. 2012;4(127):127ra37.

17. Ritprajak P, Azuma M. Intrinsic and extrinsic control of expression of the immunoregulatory molecule PD-L1 in epithelial cells and squamous cell carcinoma. Oral oncology. 2015;51(3):221-8.

18. Green MR, Monti S, Rodig SJ, Juszczynski P, Currie T, O'Donnell E, et al. Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. Blood. 2010;116(17):3268-77.

19. Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202-9.

20. Chen BJ, Chapuy B, Ouyang J, Sun HH, Roemer MG, Xu ML, et al. PD-L1 expression is characteristic of a subset of aggressive B-cell lymphomas and virus-associated malignancies. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013;19(13):3462-73.

21. Fang W, Zhang J, Hong S, Zhan J, Chen N, Qin T, et al. EBV-driven LMP1 and IFN-gamma up-regulate PD-L1 in nasopharyngeal carcinoma: Implications for oncotargeted therapy. Oncotarget. 2014;5(23):12189-202.

22. Lyford-Pike S, Peng S, Young GD, Taube JM, Westra WH, Akpeng B, et al. Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma. Cancer research. 2013;73(6):1733-41.

23. Yang W, Song Y, Lu YL, Sun JZ, Wang HW. Increased expression of programmed death (PD)-1 and its ligand PD-L1 correlates with impaired cell-mediated immunity in high-risk human papillomavirus-related cervical intraepithelial neoplasia. Immunology. 2013;139(4):513-22.

24. Bellucci R, Martin A, Bommarito D, Wang K, Hansen SH, Freeman GJ, et al. Interferon-gamma-induced activation of JAK1 and JAK2 suppresses tumor cell susceptibility to NK cells through upregulation of PD-L1 expression. Oncoimmunology. 2015;4(6):e1008824.

25. Movva S, Wen W, Chen W, Millis SZ, Gatalica Z, Reddy S, et al. Multi-platform profiling of over 2000 sarcomas: identification of biomarkers and novel therapeutic targets. Oncotarget. 2015;6(14):12234-47.

26. Velcheti V, Rimm DL, Schalper KA. Sarcomatoid lung carcinomas show high levels of programmed death ligand-1 (PD-L1). J Thorac Oncol. 2013;8(6):803-5.

27. Cao Y, Zhang L, Kamimura Y, Ritprajak P, Hashiguchi M, Hirose S, et al. B7-H1 overexpression regulates epithelial-mesenchymal transition and accelerates carcinogenesis in skin. Cancer research. 2011;71(4):1235-43.

28. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517(7536):576-82.

29. Seiwert TY, Zuo Z, Keck MK, Khattri A, Pedamallu CS, Stricker T, et al. Integrative and Comparative Genomic Analysis of HPV-Positive and

HPV-Negative Head and Neck Squamous Cell Carcinomas. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21(3):632-41.

30. Jordan RC, Lingen MW, Perez-Ordonez B, He X, Pickard R, Koluder M, et al. Validation of methods for oropharyngeal cancer HPV status determination in US cooperative group trials. The American journal of surgical pathology. 2012;36(7):945-54.

31. Dovedi SJ, Adlard AL, Lipowska-Bhalla G, McKenna C, Jones S, Cheadle EJ, et al. Acquired resistance to fractionated radiotherapy can be overcome by concurrent PD-L1 blockade. Cancer research. 2014;74(19):5458-68.

32. Zhang P, Su DM, Liang M, Fu J. Chemopreventive agents induce programmed death-1-ligand 1 (PD-L1) surface expression in breast cancer cells and promote PD-L1-mediated T cell apoptosis. Molecular immunology. 2008;45(5):1470-6.

33. Dong H, Strome SE, Salomao DR, Tamura H, Hirano F, Flies DB, et al. Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. Nature medicine. 2002;8(8):793-800.

34. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome biology. 2011;12(4):R41.

35. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013;499(7457):214-8.

36. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nature methods. 2015;12(5):453-7.

37. Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of gene expression data using BRB-ArrayTools. Cancer informatics. 2007;3:11-7.

38. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(25):14863-8.

39. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603-7.

40. Byers LA, Diao L, Wang J, Saintigny P, Girard L, Peyton M, et al. An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013;19(1):279-90.

41. Lee JS, Chu IS, Mikaelyan A, Calvisi DF, Heo J, Reddy JK, et al. Application of comparative functional genomics to identify best-fit mouse models to study human cancer. Nature genetics. 2004;36(12):1306-11.

42. Lee KW, Lee SS, Kim SB, Sohn BH, Lee HS, Jang HJ, et al. Significant association of oncogene YAP1 with poor prognosis and cetuximab resistance in colorectal cancer patients. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21(2):357-64.

43. Ramaswamy S, Tamayo P, Rifkin R, Mukherjee S, Yeang CH, Angelo M, et al. Multiclass cancer diagnosis using tumor gene expression signatures. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(26):15149-54.

44. Keck MK, Zuo Z, Khattri A, Stricker TP, Brown CD, Imanguli M, et al. Integrative Analysis of Head and Neck Cancer Identifies Two Biologically Distinct HPV and Three Non-HPV Subtypes. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21(4):870-81.

45. Gronhoj Larsen C, Gyldenlove M, Jensen DH, Therkildsen MH, Kiss K, Norrild B, et al. Correlation between human papillomavirus and p16 overexpression in oropharyngeal tumours: a systematic review. British journal of cancer. 2014;110(6):1587-94.

46. Keam B, Kim S, Ahn YO, Kim TM, Lee SH, Kim DW, et al. In Vitro Anticancer Activity of PI3K Alpha Selective Inhibitor BYL719 in Head and Neck Cancer. Anticancer research. 2015;35(1):175-82.

47. Han JJ, Kim DW, Koh J, Keam B, Kim TM, Jeon YK, et al. Change in PD-L1 Expression After Acquiring Resistance to Gefitinib in EGFR-

Mutant Non-Small-Cell Lung Cancer. Clinical lung cancer. 2016;17(4):263-70 e2.

48. Kloor M, Michel S, von Knebel Doeberitz M. Immune evasion of microsatellite unstable colorectal cancers. International journal of cancer Journal international du cancer. 2010;127(5):1001-10.

49. Tang KW, Alaei-Mahabadi B, Samuelsson T, Lindh M, Larsson E. The landscape of viral expression and host gene fusion and adaptation in human cancer. Nature communications. 2013;4:2513.

50. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. The New England journal of medicine. 2015;372(26):2521-32.

51. Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. The New England journal of medicine. 2015;373(2):123-35.

52. Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Noguchi T, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. Nature. 2014;515(7528):577-81.

53. Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. PloS one. 2007;2(8):e796.

54. Rajasagi M, Shukla SA, Fritsch EF, Keskin DB, DeLuca D, Carmona E, et al. Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. Blood. 2014;124(3):453-62.

55. Dong H, Chen X. Immunoregulatory role of B7-H1 in chronicity of inflammatory responses. Cellular & molecular immunology. 2006;3(3):179-87.

56. Marzec M, Zhang Q, Goradia A, Raghunath PN, Liu X, Paessler M, et al. Oncogenic kinase NPM/ALK induces through STAT3 expression of immunosuppressive protein CD274 (PD-L1, B7-H1). Proceedings of the National Academy of Sciences of the United States of America. 2008;105(52):20852-7.

57. Klapper JA, Downey SG, Smith FO, Yang JC, Hughes MS,

Kammula US, et al. High-dose interleukin-2 for the treatment of metastatic renal cell carcinoma : a retrospective analysis of response and survival in patients treated in the surgery branch at the National Cancer Institute between 1986 and 2006. Cancer. 2008;113(2):293-301.

58. Cancer Genome Atlas Research N. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature. 2013;499(7456):43-9.

59. Perng P, Lim M. Immunosuppressive Mechanisms of Malignant Gliomas: Parallels at Non-CNS Sites. Frontiers in oncology. 2015;5:153.

60. Sprinzl MF, Galle PR. Immune control in hepatocellular carcinoma development and progression: role of stromal cells. Seminars in liver disease. 2014;34(4):376-88.

61. Ward MJ, Thirdborough SM, Mellows T, Riley C, Harris S, Suchak K, et al. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. British journal of cancer. 2014;110(2):489-500.

62. Velcheti V, Schalper KA, Carvajal DE, Anagnostou VK, Syrigos KN, Sznol M, et al. Programmed death ligand-1 expression in non-small cell lung cancer. Laboratory investigation; a journal of technical methods and pathology. 2014;94(1):107-16.

63. Thompson RH, Gillett MD, Cheville JC, Lohse CM, Dong H, Webster WS, et al. Costimulatory B7-H1 in renal cell carcinoma patients: Indicator of tumor aggressiveness and potential therapeutic target. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(49):17174-9.

64. Hino R, Kabashima K, Kato Y, Yagi H, Nakamura M, Honjo T, et al. Tumor cell expression of programmed cell death-1 ligand 1 is a prognostic factor for malignant melanoma. Cancer. 2010;116(7):1757-66.

65. Droeser RA, Hirt C, Viehl CT, Frey DM, Nebiker C, Huber X, et al. Clinical impact of programmed cell death ligand 1 expression in colorectal cancer. European journal of cancer. 2013;49(9):2233-42.

66. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. The New England journal of medicine. 2010;363(1):24-35.

67. Chung CH, Zhang Q, Kong CS, Harris J, Fertig EJ, Harari PM, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(35):3930-8.

68. Yang MH, Chang SY, Chiou SH, Liu CJ, Chi CW, Chen PM, et al. Overexpression of NBS1 induces epithelial-mesenchymal transition and co-expression of NBS1 and Snail predicts metastasis of head and neck cancer. Oncogene. 2007;26(10):1459-67.

69. Ota K, Azuma K, Kawahara A, Hattori S, Iwama E, Tanizaki J, et al. Induction of PD-L1 Expression by the EML4-ALK Oncoprotein and Downstream Signaling Pathways in Non-Small Cell Lung Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015.

70. Ebert PJ, Cheung J, Yang Y, McNamara E, Hong R, Moskalenko M, et al. MAP Kinase Inhibition Promotes T Cell and Anti-tumor Activity in Combination with PD-L1 Checkpoint Blockade. Immunity. 2016;44(3):609-21.

71. Ferris RL, Blumenschein G, Jr., Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. The New England journal of medicine. 2016;375(19):1856-67.

72. Chow LQ, Haddad R, Gupta S, Mahipal A, Mehra R, Tahara M, et al. Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. J Clin Oncol. 2016.

초 록

서론: 대표적인 면역항암제인 PD-1/PD-L1 억제제의 임상적 효과를 예측할 수 있는 인자에 대해서 PD-L1 발현, 종양침윤림프구 존재, 그리고 암 돌연변이 개수가 알려져 있지만, 이에 대한 통합적 임상적 적용과 의미에 대해서는 연구가 부족하다. 본 연구에서는 The Cancer Genome Atlas (TCGA) 데이터베이스를 이용하여, PD-L1 발현과 종양침윤림프구를 동시에 측정하여 면역항암제의 치료 기전과 연관된 유전체의 특성을 이해하고, 이들에 대한 상위 분자생물학적 기전에 대해서 통합적으로 분석하였다.

방법: TCGA 발현체 분석 (N=9,677)을 통해서 PD-L1 과 CD8A 발현을 분석하였고, 두 유전자의 중앙값에 따라 4개의 환자군으로 분류하였다 (tumor microenvironment immune types, TMIT). 서울대학교병원 코호트에서 두 단백질에 대한 면역화학염색을 시행하였고, epithelial-mesenchymal transition (EMT)과 연관된 인자도 함께 분석하였으며, 이 현상을 5개의 두경부암 세포주에서 입증하였다.

결과: TCGA 분석결과 TMIT I은 돌연변이 개수, PD-L1 유전자증폭, 그리고 암 바이러스 감염 여부와 연관성이 있었다. CD8양성 T세포가 많이 침윤된 종양미세환경 중에서도, Regulatory T세포와 M2 macrophage가 많이 침윤되어있는 경우 유의하게 안 좋은 예후를 보였다. PD-L1 발현은 EMT 현상과 연관이 있었으며, PD-L1 발현과 EMT 현상이 같이 있는 경우에는 불량한 예후와 연관이 있었다. 서울대병원 코호트 및

두경부암 세포주에서 Cisplatin 투약을 한 경우 PD-L1 발현이 증가하는 경향을 확인하였으며, MAPK/ERK 신호전달 기전의 활성화를 동반하였다.

결론: 본 연구에서는 면역항암제가 각광받는 현 시점에서 PD-L1 발현과 함께 종양침윤림프구에 대한 분석을 통합적으로 같이 진행하는 것이 중요함을 시사한다.

주요어: 종양 미세 환경; 체계적 분석; 차세대염기서열분석; 발현체분석; 면역회피; 면역관문억제제.
학 번: 2014 - 22036